Social Science
Library

MAR 19 1954

Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
AIR FORCE PERSONNEL AND TRAINING RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *AF Personnel and Training Research Center*
W. R. Garner, *The Johns Hopkins University*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *AF Personnel and Training Research Center*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

Conditioning of Muscle Action Potential Increments Accompanying an Instructed Movement: J. B. FINK 61
Use of the Blackout in the Investigation of Temporal Discrimination in Fixed-Interval Reinforcement: C. B. FERSTER 69
Probability of Response and Intertrial Association as Functions of Monocular and Binocular Stimulation: G. COLLIER 75
Voluntary Rest Pause Behavior in a Block-Turning Task as a Function of Wrist-Cuff Weight: R. F. BOLDT AND D. S. ELLIS 84
Simple Concept Learning as a Function of Intralist Generalization: M. H. BAUM 89
Increased Pronouncing Behavior as a Factor in Serial Learning: H. G. PRICE AND D. LEWIS 95
The Effect of Pacing on the Learning of a Psychomotor Response: J. A. ADAMS 101
Experimental Manipulation of Verbal Behavior: B. D. COHEN, H. I. KALISH, J. R. THURSTON, AND E. COHEN 106
Anxiety and Stress in Learning: The Role of Intraserial Duplication: R. S. LAZARUS, J. DEESE, AND R. HAMILTON 111
Intellectual Performance as a Function of Two Types of Psychological Stress: S. F. OSLER 115
Functional Fixedness as Related to Elapsed Time and to Set: R. E. ADAMSON AND D. W. TAYLOR 122
The Relation of Anxiety to Differential Eyelid Conditioning: K. W. SPENCE AND I. E. FARBER 127
On Reading Printed Matter with Interrupted Light: V. L. SENDERS 135
Erratum 136

American Psychological Association

Vol. 47 No. 2 February 1954

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1954 the publication cost is $20 per printed page. Authors of priority publications receive no free offprints.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Air Force Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 (4, Suppl.), 389–449. In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879.

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 13, 1947.

Copyright, 1954, by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 47, No. 2 FEBRUARY, 1954

CONDITIONING OF MUSCLE ACTION POTENTIAL INCREMENTS ACCOMPANYING AN INSTRUCTED MOVEMENT [1]

JOHN B. FINK

University of Louisville

If we define *Stimulus R* as the signal to which *S* is instructed to respond as rapidly as possible, *Stimulus A* as an instructed, warning signal immediately preceding Stimulus R, and *R* as the overt response which *S* is instructed to perform, then various studies (3, 7, 9, 10, 14) suggest that the muscle action potential (MAP) increments normally preceding and accompanying R are observed, as a function of repeated trials, to creep forward into the interval between the initiation of Stimulus A and the initiation of Stimulus R. Here we have the establishment of conditioned MAP increments to a stimulus designated by instructions to *S* as antecedent.

Will an antecedent stimulus about which *S* is *not* informed acquire the property of eliciting MAP increments in this situation? Specifically, can MAP increments accompanying an instructed, overt movement be conditioned to an antecedent, noninstructed stimulus?

[1] This article is based upon a dissertation submitted to the department of psychology of Indiana University in partial fulfillment of the requirements for the Ph.D. degree. The research was performed in conjunction with Air Force Contract No. 33(038)-19630. Thanks are due Dr. R. C. Davis, principal investigator of the contract.

METHOD

Experimental design.—The experimental situation was modeled after the Pavlovian paradigm. The *S* was instructed to respond to a high-pitched signal tone with a right-hand key pressing, to a low-pitched signal tone with a left-hand key pressing. He was informed that there would be a special warning tone preceding the high-pitched, right hand signal tone; but that no warning tone would occur before the low-pitched, left hand signal tone.

In the actual operating procedure, a training trial for the experimental group, E, consisted of a low level, white noise of 2-sec. duration inserted contiguously prior to the low-pitched, left hand signal tone. Our interest lay in changes in left arm MAP magnitudes during the noise interval as a function of training trials, and as a function of extinction trials in which the noise was presented without the signal tone.

Thus, if we speak of our situation in Pavlovian terms, the left hand signal tone, about which *S* was informed, was functioning as an unconditioned stimulus (UCS); and the noise, about which *S* was not informed, was placed in the temporal role of a conditioned stimulus (CS). However, in our discussion we shall generally refer to these as instructed stimulus (IS) and noninstructed stimulus (NIS).

Ten training trials were imbedded in a series totaling 20 trials, the other 10 of which consisted of the instructed warning tone followed by the high-pitched, right hand signal tone. This 20-trial series was called the training series. An

extinction series of 10 trials followed the training series.

In order to attribute any systematic effects to the pairing of white noise and left hand response signal, a control group, C, was required, to which the identical experimental situation, except for contiguity of noise and left hand response signal, was to be presented.

Detailed trial-by-trial treatment of Groups E and C during the training series and the extinction series is presented in Table 1. In this table, H represents the high-pitched signal tone, L represents the low-pitched signal tone, W represents the warning tone preceding H, and N represents the white noise presentation preceding L. Thus N—L is read, ". . . white noise presentation followed by a low-pitched signal tone," etc. As is indicated, whereas N preceded L contiguously for Group E, N and L were presented to Group C like separate trials. The time between them was 1 min., which was the standard intertrial interval for both groups.

Apparatus.—The techniques for measuring MAP's, as well as the design and testing of equipment for recording action potentials, have been discussed by Davis (2, 4), who has also described (5) the integrators and some of the accessory apparatus used in this investigation.

In brief, pickup of MAP's was effected by the use of a small (¼-in. diameter), active, plate electrode contiguous to the muscle under investigation, and a reference electrode consisting of a felt pad, ¼ in. thick and 2 × 6 in. in area, which was saturated with saline solution and fastened in contact with the upper arm. Potentials from the electrode system were fed through a vacuum tube amplifying system, through an integrator, and into a recording system. The recording panel and pens, manufactured by Electromedical Laboratories, were fed by power amplifiers. A logarithmic transformation of the microvoltage output was effected to conform with the fact that the distribution of log microvolts gives a closer approximation to the normal distribution function than does the microvolt distribution.

The tones, produced by Hewlett-Packard oscillators, were fed into an electronic timer modified from a design published by Roush (15) and from there were led into a set of headphones worn by *S*. The experimental chamber was a darkened, electrically shielded, sound-insulated room.

Subjects.—The *S*s were 20 students who claimed normal hearing. They were paid for their services. They were distributed, by a random-numbers schedule, into two groups.

Experimental procedure.—H, L, and W were 70 db tones of 1400, 700, and 1200 cps, respectively. H and L were .5 sec. in duration, W was 2 sec. in duration. N was 44 db and of 2-sec. duration.

The 70 db levels of H, L, and W were selected arbitrarily as being well above normal threshold. For N, a level was selected by *E* which was somewhat above his lower threshold. Then five preliminary *S*s who claimed normal hearing were run through a series of 10 presentations of N which were separated by intervals varying, by random selection, from 1 to 5 sec. They were told to press a response key whenever they heard a sound. Since all *S*s reported all presentations, and gave no incorrect reports, that level (44 db) was used.

All *S*s were installed in the experimental apparatus in the same way, and all were given identical instructions. The active electrode was placed upon the dorsal crest of the left forearm in the region of the extensor digitorum muscle. The *S* was seated at the response table upon which were mounted two telegraph keys, one for each hand. An elastic band held the lower part of each forearm snugly to the table. The *S*'s middle fingers were placed on the keys, and *S* was instructed as follows:

"You are going to hear two kinds of signal tones; a high-pitched signal tone, and a low-pitched signal tone. They will sound like this. (*S* is given a sample of each tone.) When you hear the high-pitched signal tone, press the right hand key just as quickly as you can. When you hear the low-pitched signal tone, press the left hand key just as quickly as you can. This is a reaction time experiment, so I'd like you to press the correct key just as soon as the signal tone comes on."

The *S* was then given a series of tones until two successive correct responses were obtained to each signal tone. If *S* responded incorrectly, the instructions were read to him again. No *S* required more than two readings of the instructions.

Following this, *S* was instructed further:

"In this experiment we are interested in finding out whether or not reaction time to a signal tone which is preceded by a warning tone is faster than reaction time to a signal tone which is not preceded by a warning tone. To investigate this, we are going to insert a 2-sec. warning tone just before the half-second, high-pitched signal tone. The warning tone will begin 2 sec. before the high-pitched signal tone, and will end at the same instant that the high-pitched signal tone comes on. There will not be any warning tone before the half-second, low-pitched signal tone.

"You will not have any difficulty distinguishing these different tones. The warning tone will be somewhat lower in pitch than the subsequent high-pitched signal tone, but considerably higher in pitch than the low-pitched signal tone.

"Your response instructions are the same as before; that is: press the correct key just as soon

as one of the signal tones comes on. When the high-pitched signal tone comes on, press the right hand key; when the low-pitched tone comes on, press the left hand key. Remember, a 2-sec. warning tone will occur before the high-pitched signal tone; no warning tone will occur before the low-pitched signal tone."

Measurement.—Measurement of MAP's was effected for both groups on the 10 experimental training trials in the 20-trial training series, and on all the trials in the extinction series. For each of these trials, five .5-sec. measurement intervals were marked off, beginning .5 sec. before the onset of the NIS and extending to its termination. Thus, measures were obtained of potentials occurring before and during the NIS.

Since the integrators functioned to provide peaks at .1-sec. intervals, there were five peaks in each designated .5-sec. interval. The three highest of these five peaks were measured, and their mean obtained to provide a measure for each .5-sec. interval.

Reaction times in the two groups were measured from the onset of the IS. For Group C this involved using the b portions of the training trials (cf. Table 1).

Statistical treatment.—To compare MAP effects in the two groups for measures from five time intervals for the 10 training trials, an analysis of variance technique for repeated measures was used. A four-way analysis for repeated measures in two independent groups of *Ss*, extended from Edwards (11), was carried out. The variables involved were groups, *Ss*, trials, and time (i.e., successive .5-sec. measurement intervals in each trial).

Initially, a three-way analysis was done for each group to obtain a within-groups sum of squares for *Ss* × time and *Ss* × trials interactions. These were later used in getting the required error terms. Within each group, the variables were *Ss*, time, and trials.

To examine the effects within each group for the time and trials variables and their interactions, *F* ratios for these were computed for each group separately (cf. Table 2). In each case, the error term was the mean square for the interaction of *Ss* with the variable under consideration.

The groups were then combined to yield between-groups sums of squares for groups, time, trials, and their interactions. The *F* ratios were computed for the between-groups variances, using the appropriate pooled within-groups sums of squares as error terms (cf. Table 3). The error term for the groups effect was the mean square for pooled *Ss*. In every other instance, the between-groups effects of the variable under consideration were tested by the mean square derived from the pooled interactions of that variable with *Ss* within the group.

Extinction trial measures were analyzed in the same way.

Reaction time data were evaluated by means of the analysis of variance technique for repeated measures as discussed by Edwards (11).

TABLE 1

TRIAL-BY-TRIAL TREATMENT FOR GROUPS E AND C

Conditioning			
Group E	Group C	Group E	Group C
1. N—L	1a. N—	11. W—H	11. W—H
	b. —L	12. N—L	12a. N—
2. W—H	2. W—H		b. —L
3. W—H	3. W—H	13. W—H	13. W—H
4. N—L	4a. N—	14. N—L	14a. N—
	b. —L		b. —L
5. N—L	5a. N—	15. W—H	15. W—H
	b. —L	16. N—L	16a. N—
6. W—H	6. W—H		b. —L
7. N—L	7a. N—	17. N—L	17a. N—
	b. —L		b. —L
8. W—H	8. W—H	18. W—H	18. W—H
9. N—L	9a. N—	19. W—H	19. W—H
	b. —L	20. N—L	20a. N—
10. W—H	10. W—H		b. —L
Extinction			

All 10 extinction trials for Groups E and C involved presentation of N only.

RESULTS

Figures 1 and 2 show MAP measures of Groups E and C, respectively, as a function of training trials for the five .5-sec. time intervals preceding and accompanying the NIS. These figures suggest that Group E MAP measures during the training series show an incremental trend as a function of both time and trials, in contrast to the absence of such a trend, and perhaps the presence of a decremental trend, for Group C. These indications are supported by analysis of variance results in Tables 2 and 3.

Figures 3 and 4 show MAP measures of Groups E and C, respectively, as a function of extinction trials, for the five .5-sec. time intervals preceding and accompanying the NIS. These figures suggest that Group E

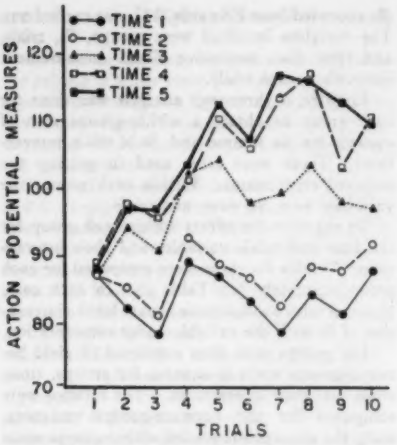


FIG. 1. Group E muscle action potential measures as a function of training trials for time intervals preceding and accompanying the non-instructed stimulus

MAP measures show a more marked decremental trend as a function of both trials and time than do the corresponding measures of Group C. Analysis of variance results in Tables 4 and 5 lend support.

Figure 5 shows the markedly different overt response time curves for

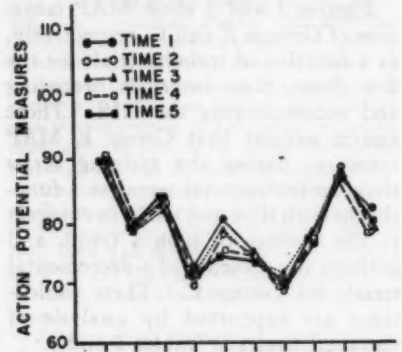


FIG. 2. Group C muscle action potential measures as a function of training trials for time intervals preceding and accompanying the non-instructed stimulus

TABLE 2

WITHIN-GROUPS ANALYSES OF VARIANCE OF MUSCLE ACTION POTENTIAL MEASURES FOR TRAINING TRIALS OF GROUPS E AND C

Group	Source of Variation	*df*	Mean Square	*F*
E	Trials	9	1,544.22	2.02*
	Ss × Trials	81	766.07	
	Time	4	12,205.75	15.93**
	Ss × Time	36	766.22	
	Time × Trials	36	315.08	3.51**
	Ss × Time × Trials	324	89.76	
	Ss	9		
	Total	499		
C	Trials	9	2,211.44	2.61*
	Ss × Trials	81	848.56	
	Time	4	28.25	1.72
	Ss × Time	36	16.44	
	Time × Trials	36	12.67	1.12
	Ss × Time × Trials	324	11.32	
	Ss	9		
	Total	499		

*p = .05.
**p = .01.

the two groups during the training trials. Analysis of variance applied to these latency measures (cf. Table 6) yielded a between-groups F ratio significant at the .05 level. The F ratios for trials and for trials × groups interaction were both significant at the .01 level. The Group E trend suggests a systematic decrement in overt re-

TABLE 3

BETWEEN-GROUPS ANALYSIS OF VARIANCE OF MUSCLE ACTION POTENTIAL MEASURES FOR TRAINING TRIALS OF GROUPS E AND C

Source of Variation	*df*	Mean Square	*F*
Groups	1	212,487.00	4.49*
Pooled *Ss*	18	47,337.77	
Trials	9	1,816.78	2.25*
Pooled *Ss* × Trials	162	807.31	
Time	4	5,865.50	14.99**
Pooled *Ss* × Time	72	391.33	
Time × Trials	36	162.50	3.22**
Pooled *Ss* × Time × Trials	648	50.54	
Time × Groups	4	6,368.50	16.27**
Trials × Groups	9	2,827.78	3.50*
Time × Trials × Groups	36	165.25	3.27**
Total	999		

*p = .05.
**p = .01.

TABLE 4

WITHIN-GROUPS ANALYSES OF VARIANCE OF MUSCLE ACTION POTENTIAL MEASURES FOR EXTINCTION TRIALS OF GROUPS E AND C

Group	Source of Variation	*df*	Mean Square	*F*
E	Trials	9	1,425.89	2.07*
	Ss × Trials	81	689.53	
	Time	4	5,293.75	12.80**
	Ss × Time	36	413.69	
	Time × Trials	36	91.58	1.67*
	Ss × Time × Trials	324	54.75	
	Ss	9		
	Total	499		
C	Trials	9	453.11	1.49
	Ss × Trials	81	303.47	
	Time	4	54.00	0.81
	Ss × Time	36	66.36	
	Time × Trials	36	10.78	1.39
	Ss × Time × Trials	324	7.74	
	Ss	9		
	Total	499		

* $p = .05$.
** $p = .01$.

sponse latencies associated with the systematic increment in MAP magnitude observed in Fig. 1.

A plot of trial means of reaction times against trial means of MAP measures suggested that a straight line function might be fitted to this relation. A least-squares fit yielded the equation

$$T = -.017P + 2.44$$

where T is reaction time in seconds, and P is MAP magnitude in log microvolts. Computation of a Pearson r for the relation yielded a value of $-.97$.

TABLE 5

BETWEEN-GROUPS ANALYSIS OF VARIANCE OF MUSCLE ACTION POTENTIAL MEASURES FOR EXTINCTION TRIALS OF GROUPS E AND C

Source of Variation	*df*	Mean Square	*F*
Groups	1	218,784.00	4.44*
Pooled *Ss*	18	49,246.28	
Trials	9	1,217.22	2.45*
Pooled *Ss* × Trials	162	496.50	
Time	4	2,152.00	8.96**
Pooled *Ss* × Time	72	240.03	
Time × Trials	36	44.39	1.42*
Pooled *Ss* × Time × Trials	648	31.24	
Time × Groups	4	3,195.75	13.31**
Trials × Groups	9	217.33	0.44
Time × Trials × Groups	36	58.47	1.87**
Total	999		

* $p = .05$.
** $p = .01$.

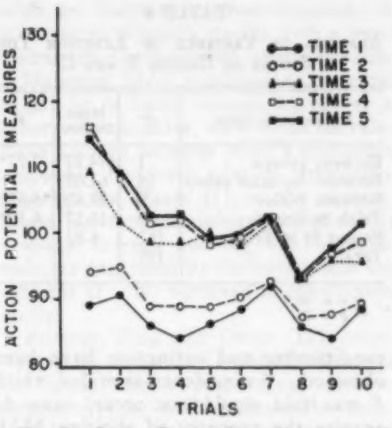


FIG. 3. Group E muscle action potential measures as a function of extinction trials for time intervals preceding and accompanying the noninstructed stimulus

DISCUSSION

It is apparent that systematic modification effects similar to those of classical

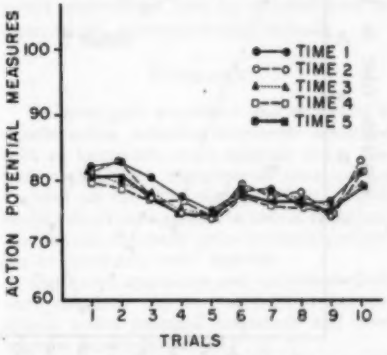


Fig. 4. Group C muscle action potential measures as a function of extinction trials for time intervals preceding and accompanying the noninstructed stimulus

TABLE 6

ANALYSIS OF VARIANCE OF REACTION TIME SCORES OF GROUPS E AND C

Source of Variation	*df*	Mean Square	*F*
Between groups	1	474.32	7.77*
Between *Ss*, same group	18	61.02	
Between trials	9	28.40	5.89**
Trials × Groups	9	21.17	4.39**
Pooled *Ss* × Trials	162	4.82	
Total	199		

*p = .05.
**p = .01.

conditioning and extinction have been obtained. An incidental stimulus, which *S* was told would not occur, came to acquire the property of eliciting MAP increments.

It is of interest to compare some of our results with those of Davis (1), although the situations are not identical. Davis' arrangement was a simple reaction time situation in which *S* was instructed to move the lever of a response key as rapidly as possible to a signal light; thus the action potentials with which Davis was concerned followed the stimulus to which *S* had been instructed to respond. In the situation of the present experiment, the measured MAP's preceded the onset of the stimulus to which *S* had been instructed to respond, but followed the onset of the antecedent stimulus which *S* had been told would not occur.

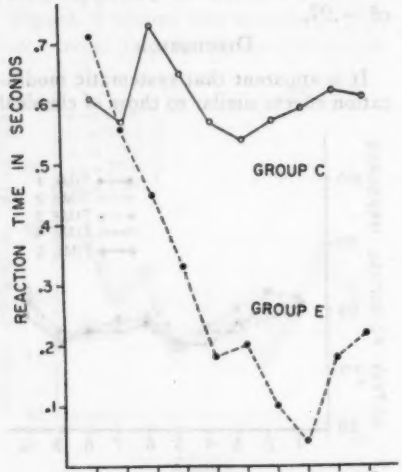


FIG. 5. Reaction times of overt response as a function of training trials

The following observations with respect to the acquired MAP increments during the antecedent stimulus interval are similar to those of Davis for a simple reaction time situation: (*a*) MAP increments begin to develop within the first .5 sec. of the interval (cf. Fig. 1 and 3). (*b*) MAP magnitude is greater at the end of the interval than at the beginning (cf. Fig. 1 and 3). (*c*) Reaction time of the subsequent overt response is a linear, inverse function of log MAP magnitude at the end of the interval. Thus it appears that, in at least these respects, acquired MAP increments resemble increments obtained via direct elicitation.

Let us consider the relation of our acquired MAP increments to classical conditioning. The UCS, as usually selected by *E*, is a stimulus which, prior to experimental operations introducing the CS, can be depended upon to elicit a defined response class. The IS, which was used to elicit the key-pressing response and its associated MAP increment, acquired this property by virtue of the specific experimental operation of *verbal instruction*. Can this signal tone now be regarded as a UCS? Perhaps so. *Unconditioned*, as operationally defined in the classical conditioning situation, means merely high elicitation strength prior to the CS–UCS contiguity situation. The operation of instruction that rendered to the IS the function of response elicitation was performed prior to the experimental operations manipulating the independent variable, noise-tone contiguity.

The *Ss* in this experiment were instructed that the signal tone for the right-hand pressing would be preceded by a warning signal. The question may be raised: Did this not predispose *Ss* to make use of *any* preliminary signal, so that what occurred with respect to the

left arm was not conditioning, but merely voluntary responding following recognition of a situation similar to that on which they had been instructed; or perhaps generalization from previous experience?

What we know of generalization of conditioning and learning in human *Ss* would indeed lead us to suspect that responding to the warning signal preceding the right hand signal tone would, via verbal mediation, increase the probability of a left-arm response to a noninstructed preliminary signal preceding the left arm signal tone. But, this does not rule out conditioning. Classical conditioning is defined in terms of systematic modification of the selected response item in the presence of a paired stimulus situation.

About our experimental situation we can say that, given organisms of the type used (human), along with their history (including verbal instructions), modification effects of the kind described were obtained. To assert that such modification effects are not conditioning would be to assert that conditioning is defined in terms of the general history of the organism, rather than in terms of characteristics of the specific S–R relation under consideration.

In general, then, we shall consider our phenomenon an instance of conditioning, although we must admit that there are features here which differ from certain other conditioning situations in the literature.

We may wonder, however, if similar MAP increments might be obtained without instructions. The studies of both Hilden (13) and Van Liere (16) suggest an affirmative answer. In neither of these studies were *Ss* instructed to respond. Although Hilden utilized MAP's associated with an overt response to shock, and Van Liere utilized MAP's to a tonal stimulus without reference to a specific overt response, both investigators reported conditioning. In the present study we have a third instance of conditioning of MAP's in human *Ss*, the important difference being that the conditioned MAP increments are associated with an instructed, overt response, in contrast to Hilden's results with a noninstructed, shock-elicited, overt response.

However, MAP conditioning, as such, might be of little interest, except as a laboratory curiosity, were it not for relations holding between MAP's and overt response. One such relation, previously reported by Davis (1), is the inverse, linear function obtaining between overt response latency and log MAP magnitude, as suggested by the correlation coefficient of −.97 obtained in the present study.

Further, Fink and Davis (12), introducing the concept of an MAP threshold for overt motor responding, demonstrated that an MAP level could be established above which, it was predicted rather precisely, the overt response would occur, and below which the overt response would not occur. Since MAP's were demonstrated to "respond," i.e., increase in magnitude, below as well as above the overt threshold, we may reasonably view overt motor responses as a consequence of MAP magnitude. Consideration of this, in conjunction with the results of Hilden, Van Liere, and the present study, suggests that MAP conditioning may underlie overt motor response conditioning.

With respect to the specific area of investigation, it seems clear that MAP increments associated with an instructed, overt movement can be conditioned to antecedent, noninstructed stimuli.

Summary

To investigate acquisition and extinction of muscle action potential increments associated with an instructed, overt response and a noninstructed CS, an experimental situation was modeled on the Pavlovian paradigm. Paired sound stimuli were utilized to obtain systematic modification of muscle action potentials preceding an instructed, overt response.

Significant acquisition and extinction effects were obtained. An inverse relation between muscle action potential magnitude and overt response latency was noted.

It was concluded that muscle action potential increments associated with an instructed, overt response can be conditioned to an antecedent, noninstructed stimulus.

With reference to an earlier study, which introduced the concept of a muscle action potential threshold for overt responding, it was suggested that overt motor response conditioning may be viewed as a consequence of muscle action potential conditioning.

References

1. Davis, R. C. Set and muscular tension. *Indiana Univ. Publ.*, Sci. Ser. No. 10, 1940.
2. Davis, R. C. Methods of measuring muscular tension. *Psychol. Bull.*, 1942, **39**, 329–346.
3. Davis, R. C. The pattern of muscular action in simple voluntary movement. *J. exp. Psychol.*, 1942, **31**, 347–366.
4. Davis, R. C. The design and testing of multiple amplifiers for action potential recording. *J. exp. Psychol.*, 1943, **32**, 270–274.
5. Davis, R. C. An integrator and accessory apparatus for recording action potentials. *Amer. J. Psychol.*, 1948, **61**, 100–104.
6. Davis, R. C. Motor effects of strong auditory stimuli. *J. exp. Psychol.*, 1948, **38**, 257–275.
7. Davis, R. C. Responses to 'meaningful' and 'meaningless' sounds. *J. exp. Psychol.*, 1948, **38**, 744–756.
8. Davis, R. C. Motor responses to auditory stimuli above and below threshold. *J. exp. Psychol.*, 1950, **40**, 107–120.
9. Davis, R. C. The stimulus trace in effectors and its relation to judgment responses. *J. exp. Psychol.*, 1952, **44**, 377–390.
10. Davis, R. C. Experimental investigation of the effect of a muscle tension reflex upon simple instructed movements. In *Electrophysiological aspects of aircraft control.* Formal Progress Report, Contract AF 33(038)–19630, Indiana Univ., Feb. 1, 1952.
11. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
12. Fink, J. B., & Davis, R. C. Generalization of a muscle action potential response to tonal duration. *J. exp. Psychol.*, 1951, **42**, 403–408.
13. Hilden, A. H. An action current study of the conditioned hand withdrawal. *Psychol. Monogr.*, 1937, **49**, No. 1 (Whole No. 217), 173–204.
14. Livingston, W. A. Action potential measurements from the arms in the foreperiod of reaction time to visual stimuli. Unpublished doctor's dissertation, Indiana Univ., 1949.
15. Roush, R. G. Tone burst generator. *Electronics*, 1947, **20**, 92–96.
16. Van Liere, D. W. Characteristics of the muscle tension response to paired tones. *J. exp. Psychol.*, 1953, **46**, 319–324.

(Received June 8, 1953)

Journal of Experimental Psychology
Vol. 47, No. 2, 1954

USE OF THE BLACKOUT IN THE INVESTIGATION OF TEMPORAL DISCRIMINATION IN FIXED-INTERVAL REINFORCEMENT[1]

CHARLES B. FERSTER

Harvard University

This paper reports an investigation of temporal discrimination in behavior under fixed-interval schedules of reinforcement.[2] It has been found that under these schedules, the rate of responding following reinforcement falls to zero and then increases to a stable value (3, 4). Skinner maintained that as a result of nonreinforcement on the occasion of recent eating the rate of responding in the first part of the fixed interval is inversely proportional to the length of time that has elapsed since the occurrence of the reinforcing stimulus (3, Ch. 7). A correlation is also to be found between the number of responses emitted in each interval and the occurrence of the reinforcement; for although the fixed-interval schedule of reinforcement is programmed by time, the number of responses which occur in each interval is found to be roughly the same. By means of a "time-out of the experiment" technique, the following experiments demonstrate the temporal properties of the reinforcing stimulus in short fixed-interval schedules and the temporal properties of the "number of responses preceding reinforcement" in long fixed-interval schedules. The change in rate of responding during a given interval may be said to be under the control of the temporal properties of a stimulus if the change in rate is eliminated when the stimulus is followed by a period of time during which the response cannot be made. The time-out period necessary to eliminate the temporal discrimination must equal, or be longer than, the period during which the stimulus would otherwise have been effective in controlling the frequency of the response.

Since it has been found that without special training pigeons do not peck in the dark, we chose the blackout as our time-out technique for demonstrating the temporal properties of the reinforcing stimulus and the temporal properties of the number of pecks preceding reinforcement (1). Thus, if a blackout period (equal to, or greater than, the duration of the temporal discrimination) following the stimulus eliminates the rate change that follows the stimulus, the change in the rate of responding can then be considered a function of the temporal properties of the stimulus.

EXPERIMENT I

Purpose.—An investigation of whether the low rate of responding following the reinforcement under a short fixed-interval schedule of reinforcement is under the control of the reinforcing stimulus.

Method.—The *S*s were two adult pigeons that had been trained for a previous experiment under a variable-interval schedule of reinforcement with 1-min. blackouts occurring either after the

[1] This experiment is part of a research program carried out under Contract N5ori-07631 between Harvard University and the Office of Naval Research, U. S. Navy (Project NR 143-943, report PP-3), under the direction of B. F. Skinner. These experiments were carried out in collaboration with B. F. Skinner. Thanks are due to R. J. Herrnstein and W. H. Morse for preparing the tables.

[2] These schedules were called periodic reconditioning by Skinner (3).

reinforcement, as a result of a peck, or independently of pecks and reinforcements (1, Exp. III). The experimental cage was an aluminum icebox which was lightproof and attenuated outside noises about 20 db. An overhead light, consisting of two 6-w. lamps, provided general illumination. The front panel of the box contained a grain magazine and an illuminated 1-in. circular hole behind which was the key, a hinged Plexiglas assembly. Pecking the key operated an electrical switch which permitted automatic recording and programming of the experiment. Reinforcements were programmed by presenting an illuminated hopper of grain to *S* for 4 sec.[2]

The *S*s were placed on a 90-sec. fixed-interval schedule. After approximately 1,200 reinforcements, when a pause after the reinforcement appeared consistently, a 30-sec. blackout was introduced after every tenth reinforcement. The nine intervals without blackout, preceding the blackout interval, provided a baseline with a single *S* and a single day from which to judge the effect of the blackout on the pause at the start of the interval.

The pecks were recorded on a cumulative graphic recorder from which it was possible to determine the number of pecks to an accuracy of 1 peck when the number of responses was under 5 and to an accuracy of 5 pecks for larger numbers. The blackout was programmed by turning out both the overhead and key lights. For the duration of the blackout, the reinforcement clock was stopped and the paper-feed on the recorder disconnected. The key was still connected to the recorder, however, so that any pecks occurring in the dark would have been recorded as a vertical line on the cumulative record. This procedure of alternating nine reinforcements without blackout with one reinforcement with blackout was continued for six days with approximately 35 reinforcements each day.

Results.—The time between the reinforcement and the first response was measured to the nearest 2 sec. with a magnifying comparator on the cumulative graphic record. Whenever the reinforcement was followed by a blackout, the pause at the start of the interval was markedly decreased. These data are summarized in Table 1, which compares the time from the start of the interval to the first response for blackout and no-blackout conditions.

[2] A detailed account of the methods used in this research may be found in (2).

TABLE 1

LENGTH OF PAUSE AFTER REINFORCEMENT UNDER 90-SEC. FIXED-INTERVAL SCHEDULE OF REINFORCEMENT

(Every tenth reinforcement is followed by a 30-sec. blackout.)

S	Condition	Time to First Response (in Sec.)			
		Q_1	Median	Q_3	*N*
S_1	No blackout	14	20	22	203
	Blackout	0	2	4	21
S_2	No blackout	18	28	36	221
	Blackout	0	2	2	25

The 25th and 75th percentiles show no overlap for either *S* and indicate the order of magnitude of the effect of the blackout.

Figure 1 is a graphic cumulative record showing an interval in which the pause at the start of the interval has been eliminated by the blackout.

Discussion.—By showing that the blackout decreases or eliminates the pause at the start of the interval, we have demonstrated that the reinforcing stimulus has temporal properties. The lower rate of responding in the early part of the 90-sec. fixed interval is a temporal discrimination, under the control of the reinforcing stimulus. The blackout prevents *S* from pecking during the period in which the reinforcement would otherwise control a low rate of responding.

This experiment demonstrates that the reinforcing stimulus is effective for a period equal to or less than the duration of the blackout. The temporal properties of the reinforcement could be measured more accurately by varying the length of the blackout and observing the change in the pause at the start of the interval.

EXPERIMENT IIa

Purpose.—An investigation to determine the carry-over, from one interval to the next, of the number of responses emitted during the interval

under long fixed-interval schedules of reinforcement.

Method.—The *S*s were two adult pigeons that had extensive reinforcement on 10-, 20-, and 30-min. fixed-interval schedules. The apparatus, blackout procedure, and recording technique were the same as those used in Exp. I. The *S*s were placed on a 45-min. fixed-interval schedule of reinforcement. S_3 was placed in the apparatus in the morning and remained there until 5 P.M., and S_4 was placed in the apparatus at 5 P.M. and remained there until morning. Five-hundred reinforcements were received by S_3 under this schedule and 900 by S_4. This amount of reinforcement produced an end state of the 45-min. fixed-interval schedule. The procedure was then changed to 45-min. fixed-interval reinforcement with a 20-min. blackout after each reinforcement.

Results.—Measurements were made of the time to the first response in the interval and the number of responses in the first 5 min. of the interval. These data were then compared for the 6 days preceding and the 6 days following the introduction of the blackout. Table 2 presents a summary of these data. The blackout had two effects on the pigeons' behavior: (*a*) the first response in the interval was postponed, and (*b*) the number of responses emitted during the first 5 min. of the interval was decreased. As in Exp. I, the 25th and 75th percentiles under the two conditions show no overlap for either pigeon and indicate the large magnitude of the effect.

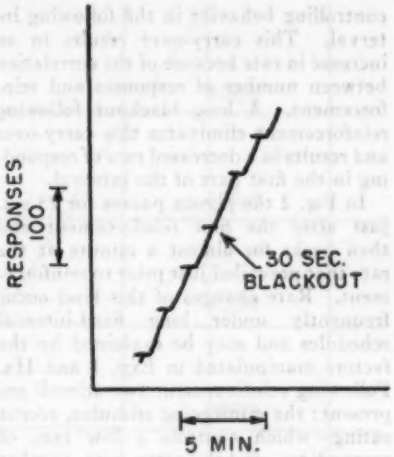


FIG. 1. A cumulative response curve of a pigeon under a 90-sec. fixed-interval schedule of reinforcement. At the arrow a 30-sec. blackout occurs following the reinforcement and eliminates the pause following the reinforcement.

The introduction of the blackout also resulted in an increase in the orderliness of the rate changes during the interval, as shown in Fig. 2 and 3, cumulative graphic curves obtained with and without blackout, respectively.

Discussion.—Experiment IIa demonstrates that under a 45-min. fixed-interval schedule of reinforcement, the number of pecks emitted prior to the reinforcement is a stimulus with temporal properties,

TABLE 2

LENGTH OF PAUSE AFTER REINFORCEMENT AND NUMBER OF RESPONSES EMITTED DURING THE FIRST 5 MIN. OF THE INTERVAL UNDER 45-MIN. FIXED-INTERVAL REINFORCEMENT

(Six days of no blackout after reinforcement were followed by six days of 20-min. blackout after each reinforcement. The first interval of each daily session was not included.)

S	Condition	Time to First Response (in Sec.)			Number of Responses in First 5 Min. of Interval			
		Q_1	Median	Q_3	Q_1	Median	Q_3	*N*
S_3	No blackout	6	30	75	4	12	18	42
	Blackout	93	380	870	0	1	6	
S_4	No blackout	0	0	1	25	50	95	24
	Blackout	120	260	480	0	5	10	

controlling behavior in the following interval. This carry-over results in an increase in rate because of the correlation between number of responses and reinforcement. A long blackout following reinforcement eliminates this carry-over and results in a decreased rate of responding in the first part of the interval.

In Fig. 2 the pigeon pauses for 25 sec. just after the first reinforcement and then pecks for almost a minute at the rate that prevailed just prior to reinforcement. Rate changes of this kind occur frequently under long fixed-interval schedules and may be explained by the factors manipulated in Exp. I and IIa. Following reinforcement two stimuli are present: the reinforcing stimulus, recent eating—which controls a low rate of responding—and the carry-over, number of pecks in the preceding interval—which controls a high rate. The 25-sec. pause, which occurs just after reinforcement, is under the control of recent eating; and the return to the rate prevailing at the time of reinforcement is under the control of the number of responses in the preceding interval.

Experiment IIb

Purpose.—An investigation parallel to Exp. IIa, in which blackouts following alternate reinforcements replace successive blocks of no-blackout/blackout conditions of the earlier experiment.

Method.—The *S*s were the two pigeons used in Exp. IIa. The apparatus, recording technique, and schedule of reinforcement were the same as those in Exp. IIa. Twenty-minute blackouts followed alternate reinforcements.

Results.—Table 3 summarizes measurements of the time from the reinforcement to the first response and the number of responses during the first 5 min. of the interval for the 6 days of the experiment. The effects of the blackout on the first part of the interval were the same as in Exp. IIa: The

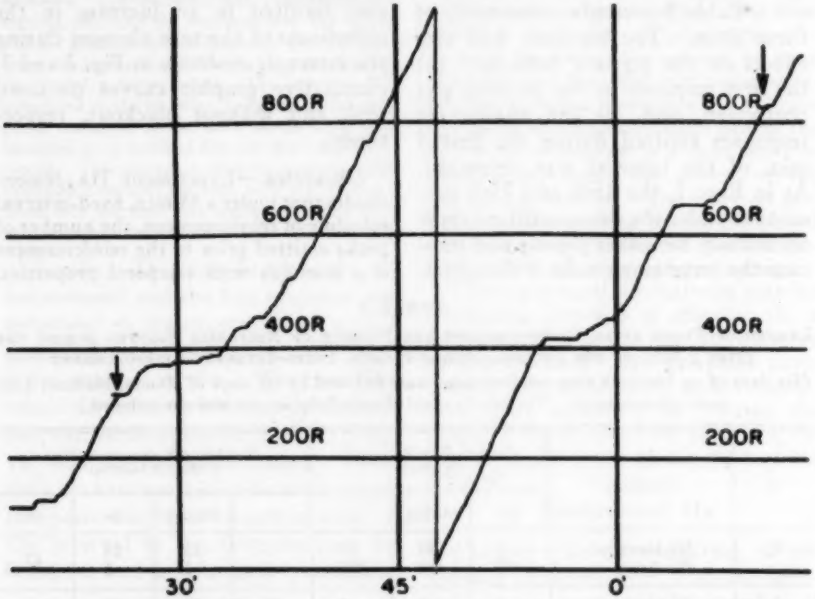


Fig. 2. A cumulative response curve of a pigeon under a 45-min. fixed-interval schedule of reinforcement. The arrows point to the reinforcements.

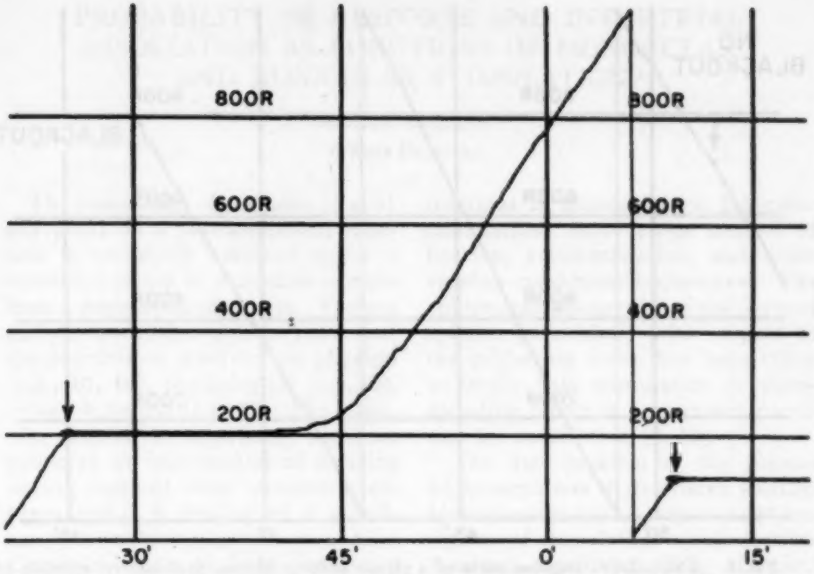


FIG. 3. A cumulative response curve of a pigeon under a 45-min. fixed-interval schedule of reinforcement with a 20-min. blackout following each reinforcement. The blackout time is not included in the record. The arrows point to reinforcements.

first response was postponed, and the number of responses emitted was decreased. As in the previous experiments, the 25th and 75th percentiles under the two conditions show no overlap for either bird and indicate the large magnitude of the effect of the blackout.

Figure 4 is a graphic cumulative record which shows one interval with no blackout following the reinforcement and a second interval with blackout. The rate of responding throughout the first interval is the same as the rate that prevailed just prior to the reinforcement. Sustained pecking throughout the no-blackout interval, as shown in Fig. 4, occurs frequently

TABLE 3

LENGTH OF PAUSE AFTER REINFORCEMENT AND NUMBER OF RESPONSES EMITTED DURING THE FIRST 5 MIN. OF THE INTERVAL UNDER 45-MIN. FIXED-INTERVAL REINFORCEMENT

(Six days of 20-min. blackouts following alternate reinforcements, with the first interval of each daily session not included)

S	Condition	Time to First Response (in Sec.)			Number of Responses in First 5 Min. of Interval			
		Q_1	Median	Q_3	Q_1	Median	Q_3	N
S_3	No blackout	0	0	95	10	127	263	15
	Blackout	487	720	1,593	0	0	0	
S_4	No blackout	0	0	10	20	95	150	18
	Blackout	157	212	360	0	5	10	

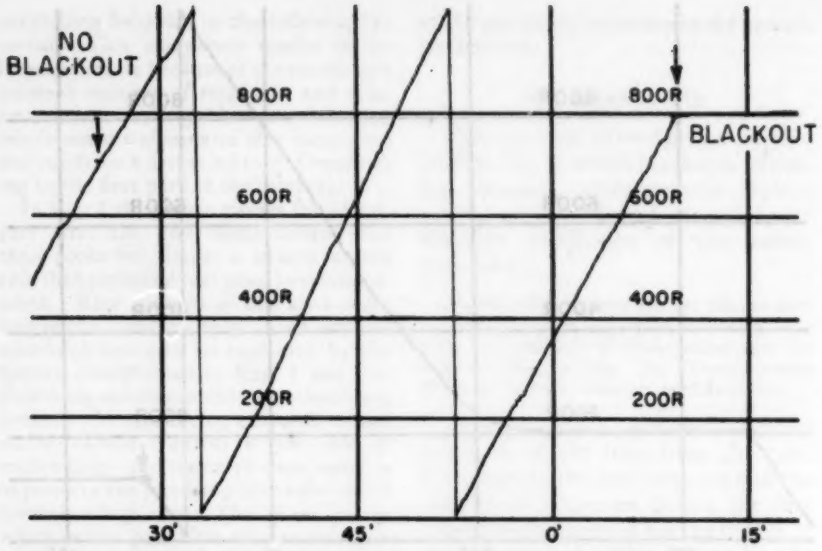


FIG. 4. A cumulative response curve of a pigeon under a 45-min. fixed-interval schedule of reinforcement with a 20-min. blackout following alternate reinforcements. The blackout time is not included in the record. The arrows point to the reinforcements.

under the alternate-blackout procedure.

Discussion.—Although the no-blackout intervals which occurred in a block (Exp. IIa) and the no-blackout intervals which were alternated with blackout intervals (Exp. IIb) are both influenced by behavior in the preceding interval, the stimulus that is carried over is not necessarily the same. The correlation between number of responses and reinforcement is clearer when blackouts occur between intervals. This correlation probably increases the influence of the stimulus on the interval that follows; e.g., if the number of blackout intervals is increased, there will be an increase in the number of responses emitted in the intervals without blackouts.

SUMMARY

Three experiments were reported in which blackouts were used to demonstrate that rate changes in the pecking behavior of pigeons on fixed-interval schedules of reinforcement were under the control of stimuli with temporal properties.

Under a 90-sec. fixed-interval schedule, the low rate of responding in the first part of the interval was under the control of the reinforcement, which had an effective duration of 30 sec. or less.

Under a 45-min. fixed-interval schedule, the rate of responding in the first part of the interval was under the control of the number of responses emitted during the preceding interval. A blackout following the reinforcement prevented the carry-over of the stimulus from one interval to the next.

REFERENCES

1. FERSTER, C. B. Sustained behavior under delayed reinforcement. *J. exp. Psychol.*, 1952, **4**, 27–45.
2. FERSTER, C. B. The use of the free operant in the analysis of behavior. *Psychol. Bull.*, 1953, **50**, 263–274.
3. SKINNER, B. F. *The behavior of organisms.* New York: D. Appleton-Century, 1938.
4. SKINNER, B. F. Some contributions of an experimental analysis of behavior to psychology as a whole. *Amer. Psychologist*, 1953, **8**, 69–78.

(Received April 6, 1953)

Journal of Experimental Psychology
Vol. 47, No. 2, 1954

PROBABILITY OF RESPONSE AND INTERTRIAL ASSOCIATION AS FUNCTIONS OF MONOCULAR AND BINOCULAR STIMULATION[1]

GEORGE COLLIER[2]
Indiana University

The probability of response (P_R) at any point on a psychophysical function is commonly assumed to be a statistical result of a random sample from a population of events. Various authors have attempted to show that the populations involved are physical (e.g., **10, 14**), physiological (e.g., **5**), or psychological (e.g., **9**). The binocular summation experiment has been proposed as one method of deciding among some of these alternative explanations. A finding of a significantly lower binocular than monocular threshold implies that the population of events involved cannot be exclusively retinal in origin and excludes, therefore, an interpretation of the P_R function entirely in terms of retinal or physical events.

The findings with respect to the occurrence of binocular summation have not been unequivocal, both results having been found with about equal frequency (e.g., **1, 3, 4, 5, 6, 7, 13, 15, 16**). These divergent results have been accounted for, in one direction or the other, in terms of the lack of control of the appropriate extraneous variables. For example, the retinal gradient of sensitivity and the resultant lack of homogeneity for different areas of the retina, along with the differential pupillary and fixation reactions to monocular vs. binocular stimulation, make strict control of fixation, accommodation, and other viewing conditions imperative. The anatomical topography of the nervous system, particularly with respect to the projection areas, has been taken to imply that stimulation of corresponding points is a necessary condition for summation (**4, 16**).

The first problem of the present experiment was to determine whether binocular summation occurs under conditions of stimulation which produce maximal sensitivity (**10**). Further, the effect of controlled variation of fixation and viewing conditions was to be determined.

A second assumption commonly made with respect to the P_R is that the population of events being sampled has fixed parameters or parameters varying randomly in time. Recent experiments (**17, 18**) have shown that responses to a repeated, invariant, simple binocular visual stimulus violate this assumption. The probability of a "yes" (or a "no") response on any trial is associated with the response (or responses) on the preceding trial (or trials) resulting in runs of "yes" and "no" responses. Thus, the observed P_R for a given intensity of stimulation is a sample from a population with nonrandomly varying parameters. No account of the associative effect is available at present. It is possible that the effect may be the result of such factors as nonrandom variations in fixation or pupillary size or that it may be the result of some nonperipheral process. The problem

[1] The experiment reported in this paper is a portion of a thesis submitted to the Graduate School of Indiana University in partial fulfillment of the requirements for the Ph.D. degree. It was supported by ONR Contract N6onr-180, T. O. IV. The writer is indebted to Dr. William S. Verplanck and Dr. William K. Estes who successively directed the dissertation.

[2] Now at Duke University.

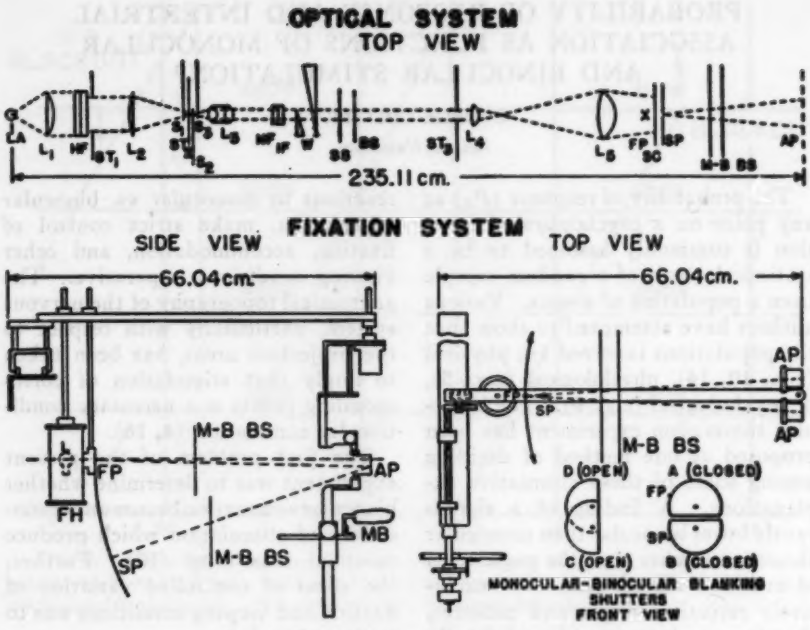


FIG. 1. Schematic diagram of the apparatus

of the retinal or nonretinal locus of the events involved in intertrial association is similar to the problem of summation for P_R. Again, it should be possible to answer the question of the retinal or nonretinal origin of the associative effect by an examination of the augmentation, reduction, or lack of change in the number of runs under monocular stimulation as compared with binocular. The problems of control of fixation, accommodation, and other viewing conditions are similar to those in the summation problem.

The second problem of this experiment, then, was a determination of the retinal vs. the nonretinal origin of the associative effect, leading to a delimitation of the possible classes of variables to be examined in attempting to account for the associative effect.

METHOD

Apparatus.—A schematic diagram of the apparatus is shown in Fig. 1.[3] The apparatus was designed to present S with short (.001 sec.) peripheral (20°), monochromatic (510 mμ max. transmission, half width 9 mμ), circular flashes of light of small visual angle (10′) at a constant intertrial interval (4.67 sec.). The brightness of the stimulus patch was continuously variable over a wide range of values. An image of the filament of a strip filament, 6 v., 18 amp. lamp (LA) provided the illumination for the stimulus patch. The system was photometrically recalibrated whenever a current decrement of .05 amp. in the lamp current occurred. Condensing lenses (L_1 and L_2) brought the image of the filament to focus at the 1-mm. strip stop (ST_2) located immediately before the fast blade of a three-bladed rotating shutter (S_2). The image of this stop was projected in a collimated beam through the filter and wedge assembly (NF, IF, and W) and was projected onto a ground glass screen (SC). The visual angle of the stimulus

[3] This apparatus was constructed by the author for use on ONR Contract N6onr-180, T. O. IV.

patch was 10′ produced by a circular aperture in a shim metal stop (*SP*) placed immediately in front of the ground-glass screen which was located in a light-tight, ventilated, sound-deadened, $3\frac{1}{2} \times 3\frac{1}{2} \times 7$ ft. chamber for *S*. An infrared filter (*HF*) and a stop (ST_1) were located between L_1 and L_2.

The distances from the artificial pupils to the stimulus patch (*SP*) and to the fixation point (*FP*) were both 50 cm. The fixation point was red and subtended 10′ of visual angle. The stimulus patch appeared 20° below the fixation point. The artificial pupils were 3 mm. in diameter. The position of the *S*'s eyes with respect to the artificial pupils was fixed by the use of a mouth bite (*MB*). The interpupillary distance of the artificial pupils and the position of the mouth bite with respect to the artificial pupils was adjustable in the vertical and horizontal dimensions, and reproducible to .5 mm.

Four blanking shutters (*M–B BS*), separately operated by Telechron motors, were mounted between the artificial pupils and the stimulus patch. These shutters were so located that it was possible to interrupt the visual path between either eye and the stimulus patch and the fixation point separately or in any combination. The operation of these shutters was faintly audible to some *S*s.

The *S* responded, if he saw a flash of light, by means of a telegraph key. The shutter was operated continuously and provided no cue as to the occurrence or nonoccurrence of a flash. Thus, only "yes" responses to a "seen" flash were possible.

Calibrations.—The flash duration of the shutter was calibrated with a photoelectric cell, an oscilloscope, and a Hewlett-Packard audio-oscillator. In seconds per degree of aperture, the flash duration was 1.852×10^{-4} sec. per degree with a pooled standard deviation (for all determinations) of 4.9×10^{-6} sec. per degree. Thus, for the aperture setting used, 5.4°, the mean flash duration was .001 sec. with a standard deviation of .00003 sec.

The intertrial interval was calibrated with a photoelectric cell and an electric timer. The mean value for the setting used was 4.67 sec. with a sigma of .03 sec.

The transmission characteristics of the interference filter (*IF*), Baird-Associates No. 7-2102-A, was determined with a Beckman spectrophotometer. It had a peak transmission of 510 mμ with a half width of 9 mμ.

The brightness of the stimulus patch was calibrated with a Macbeth illuminometer. The brightness of the patch was determined for six wedge settings. The average standard deviation for a typical set of determinations was .03 log micromicrolamberts (μμl.) The step interval on the wedge scale was .01 log μμl.

Subjects.—Nine highly trained *S*s, four women and five men were used throughout all experiments. All *S*s were paid or were members of the project, and met or exceeded service norms, Naval visual standards by the Orthorater (20/20 visual acuity, O. D., O. S.; no phorias; "color-normal"). All *S*s had previous experience with visual experimentation on the Hecht-Shlaer adaptometer and had just completed 14 days of experimentation on the present apparatus. They were not informed of the purpose, procedure, or design of the experiment. They were instructed to maintain fixation and to report "yes" by pressing the response key whenever they saw the flash. When each session began, each *S* had spent 30 min. in red goggles and 10 min. in complete darkness. All *S*s were run at the same time each day.

Procedure.—Four experiments, each involving several daily sessions, were performed. Each session consisted of six blocks of 50 stimulations at a single luminance following an initial set of "warm-up" stimulations. In each of the four experiments, three combinations of fixation and viewing conditions were used and assigned in a predetermined random order to two of the blocks of stimulations. A rest period of 1 to 2 min. was given at the end of each 100 trials, so that the fixation and viewing conditions changed three times during the course of stimulation and three times during a rest period.

Three fixation conditions were possible: binocular fixation (BF), monocular fixation right eye (RF), and monocular fixation left eye (LF), in which *S* saw the *fixation point* with both eyes, the right eye, or the left eye, respectively. Three viewing conditions were also possible: binocular viewing (BV), monocular viewing right eye (RV), and monocular viewing left eye (LV), in which *S* saw the *stimulus patch* with both eyes, the right eye, or the left eye, respectively. Seven combinations of the three fixation and the three viewing conditions were used and combined into three experiments. In Exp. 1, Cond. BF and BV, RF and RV, and LF and LV were used; in Exp. 2, Cond. BF and BV, BF and RV, and BF and LV were used; and in Exp. 3, Cond. BF and BV, RF and BV, and LF and BV were used. In Exp. 4, the same fixation and viewing conditions were used as in Exp. 1, but without the artificial pupils. Experiment 1 was run for four sessions; Exp. 2 and 3 for two sessions each; and Exp. 4 for one session for seven *S*s and three days for two *S*s.

Experiment 1 involved the conventional technique of concomitant variation of fixation and viewing conditions. In Exp. 2, fixation was always binocular and viewing conditions varied,

which should produce optimal fixation under all conditions. In Exp. 3, viewing was always binocular and fixation was varied, which should produce poor fixation, since, under the monocular fixation conditions, one eye was stimulated but not fixated. Experiment 4, when compared with Exp. 1, allows for an evaluation of the effect of artificial pupils.

In each of the experiments a single luminance method of stimulus presentation was used which allowed for a large number of trials to be used in estimating the P_R and made possible the calculation of a measure of association. The luminance for each S was chosen to yield approximately 75% "yes" responses binocularly in the first, second, and fourth experiments, and 50% "yes" responses in the third experiment. This value was determined at the beginning of each session and was held constant throughout a session.

Results

The basic data of this experiment are measures of probability of response and intertrial association, and will be presented in that order.

Probability of response.—The mean percentages of "yeses" averaged over blocks, sessions, and *S*s for all experiments are shown in Table 1. In Exp. 1, in which the fixation and viewing conditions were varied concomitantly, the percentage of "yeses" under monocular conditions was approximately one-half of the percentage under the binocular condition. The expected value for the binocular percentage "yeses," calculated on the assumption of independent but not mutually exclusive events ($P_B = P_R + P_L - P_R P_L$) from the monocular values, is 59.9%. The discrepancy between the obtained (75.0%) and expected value is 15.1%.[4] When the expected values are calculated for each *S* on each day, the obtained binocular value was above the expected value in 34 out of the 36 cases (9 *S*s by 4 days). It was above the largest monocular value in 35 out of 36 possible comparisons. Thus, summation is significantly ($p < .01$) present.

TABLE 1

Monocular and Binocular Mean Percentage of "Yeses"

Exp.	Condition		Percentage "Yeses"		
	Fixation	Viewing	Mean	Expected	*SD*
1	R	R	36.1		11.0
	L	L	37.2		12.7
	B	B	75.0	59.9	9.5
2	B	R	34.0		16.9
	B	L	38.1		16.1
	B	B	72.3	59.2	11.7
3	R	B	41.8		10.2
	L	B	44.6		10.5
	B	B	56.6		11.9
4	R	R	40.4		9.1
	L	L	36.9		11.0
	B	B	78.9	62.4	13.5

The mean percentages of "yeses" in Exp. 2, in which the fixation condition was always binocular and the viewing conditions varied, are very similar to those obtained in Exp. 1. In this case the discrepancy between the obtained (72.3%) and expected value (59.2%) for the binocular condition was 13.1%. Seventeen out of the 18 possible (9 *S*s by 2 days) obtained binocular values are above the expected value. Eighteen out of 18 are above the largest monocular value. Summation is again significantly present, and the similarity of the results with those of Exp. 1 indicate that the potential improvement of fixation resulting from having both eyes fixated under all conditions did not occur.

In the case of Exp. 3, in which the fixation conditions were varied and the viewing condition was always binocular, monocular percentages of "yeses" were approximately 80% of the binocular percentages. Sixteen of the 18 possible (9 *S*s by 2 days) binocular (fixation) values are above

[4] Chi square cannot be used to evaluate the significance of the difference since the P_R (R) and P_R (L) are empirical values. The calculated P_R (B) is subject to sampling error.

the highest monocular value. Thus, binocular (viewing) percentage of "yeses" with only one eye fixated is significantly less than the binocular percentage of "yeses" with both eyes fixated.

In Exp. 4, in which the fixation and viewing conditions were the same as Exp. 1, but in which no artificial pupils were used, the percentage of "yeses" under the monocular condition was again approximately one-half of the percentage under the binocular condition. The discrepancy between the obtained (78.9%) and the expected (62.4%) binocular value was 16.5%. Twelve out of the 13 possible (7 *Ss* by 1 day; 2 *Ss* by 3 days) binocular values are above the expected value.

The possibility that the significant amount of summation reported in Exp. 1, 2, and 4 was an artifact of the combination of independent nonrandom sequences, i.e., the significant intertrial association reported in the next section, was checked by an empirical sampling experiment. Independent nonrandom sequences of pluses and minuses were constructed by a semirandom rearrangement from a table of random numbers and combined two by two. The resultant value was compared with the value expected by the law for combination of independent nonmutually exclusive probabilities. The maximum significant increment in the obtained value above the expected value was found when the pluses greatly exceeded the minuses and amounted to approximately 1%. This value differs significantly from the mean of 14.9% obtained in the experiment, leading to the rejection of the hypothesis that the summation observed is an artifact of the nonrandomness of response.

A pooled test of the homogeneity of the variances of the mean percentage of "yeses" for right, left, and both eyes under all conditions did not lead to rejection of homogeneity ($F = .58$; $df = 11, 6373$; $p > .05$). Thus, no one of the conditions tends, systematically, to produce greater variability in percentage of "yeses."

Inter-experiment comparisons between percentage of "yeses" in the three experiments are not possible since different luminances were used in each session. Only one subject showed consistent differences between the monocular right and monocular left percentage of "yeses." The difference between the expected and observed values in Exp. 1 and 2 was not influenced by the size of the differences in P_R between the eyes. One quite interesting result occurred in three sessions for two *Ss* in which one of the monocular P_R's was zero, but in which the binocular P_R showed a significant amount of summation. This indicates the possibility of subliminal summation.

Intertrial association.—The intertrial association data were analyzed by a nonparametric serial correlation test (**12**, p. 182–183) where the amount of correlation was determined from the number of runs occurring in a sequence of responses. A "run" was defined as a sequence of responses of the same sort bounded by responses of the other kind. The test involves a comparison of the observed serial correlation (R) with an expected correlation $[E(R)]$. The hypothesis of no association can be tested by means of a critical ratio $CR = R - E(R)/\sigma_R$.[5] A CR was calculated for each S on each block of trials. The $R - E(R)$ values for Exp. 1 plotted against P_R are presented in Fig. 2. The confidence belts are the

[5] $R = 50 - 4(r)$; $E(R) = (S_1^2 - 50)/49$; $\sigma^2_R = 50 + (S_1^4 - 1965S_1^2 + 2400)/2352 - E^2(R)$; r = number of runs; S_1 = number of "yeses" minus number of "noes."

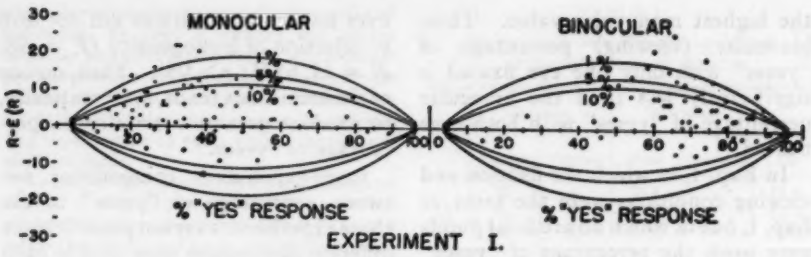


FIG. 2. Distribution of $R - E(R)$ differences in Exp. 1

significance levels for the *CR*. Two things can be noted from this plot. First, the number and magnitude of the *CR*'s do not appear to be functions of P_R. Second, very few of the values reach the 5% level of confidence, but approximately 75% of the values are greater than zero. These results are typical of those for Exp. 2, 3, and 4. A χ^2 technique was used to pool the data (**12**, Theorem I, p. 136, Theorem III, p. 138). The results of this analysis are presented in Table 3. In all but three cases (Exp. 2—BF, LV and BF, BV; Exp. 4—RF, RV), the test rejected the hypothesis of no association.

TABLE 2

MONOCULAR AND BINOCULAR MEAN χ^2

Exp.	Condition		Mean χ^2	
	Fixation	Viewing	*df**	χ^2
1	R	R	69	179.5**
	L	L	65	127.7**
	B	B	67	153.9**
2	B	R	34	84.5**
	B	L	35	38.7
	B	B	34	38.9
3	R	B	36	65.9**
	L	B	36	80.0**
	B	B	36	112.8**
4	R	R	18	26.6
	L	L	18	30.5**
	B	B	17	51.0**

* Cases in which percentage of "yeses" was so low as to invalidate the runs test were not included.
** Significant at 1% level.

The reduction in the level of significance for the *CR*'s from that observed in previous experimentation on intertrial association (**17, 18**) is probably a result of the smaller blocks of trials used; $N = 50$ vs. $N = 300$. A test of this hypothesis was made by means of a Festinger test in which blocks of trials sampled from a previous experiment were compared with the present data. The hypothesis was not rejected.

A significant ($p < .01$) decrement in the P_R between successive blocks of trials under the same condition within a session prevented a pooling of the raw data across blocks in order to increase the sample size. This decrement was atypical when compared with previous results (**18**).

Intra-experiment comparisons.—The difference between conditions within each experiment was tested by the use of a direct difference *t* test between the mean *CR*'s, since each *S* appeared under each condition. Of the 12 comparisons made, only the difference between the BF, BV condition and LF, BV condition in Exp. 3 proved to be significant at the 5% level.

Inter-experiment comparisons.—The inter-experiment comparisons were made with the same test used in the intra-experiment comparisons. Of the 12 comparisons made between the comparable conditions in each experiment, only the difference between the

RF, RV condition in Exp. 1 and the RF, RV condition in Exp. 4 proved to be significant at the 5% level.

Discussion

Probability of response.—Experiments 1, 2, and 4 show, first, that the P_R is significantly greater binocularly than monocularly for all *S*s; second, that the amount of summation is significantly greater than that which would be expected by the combination of independent but nonmutually exclusive probabilities; and third, that these results are not artifacts of the nonrandomness of the successive responses.

In order to interpret these findings, the peripheral stimulating conditions which have been suggested as variables governing the occurrence or nonoccurrence of binocular summation must be examined.

Among the variables suggested as important by various authors have been the following:

1. Graham (6) and others argue that differential pupillary reactions under monocular and binocular stimulation may result in artifactual summation. The results of Exp. 1 and 4, with and without artificial pupils, indicate that this is not an important variable in the present experiment.

2. Crozier (4) has argued that, in order to obtain summation, it is necessary to insure ease of accommodation by having the fixation point and the stimulus patch at the same viewing distance. Hewson (11), on the other hand, argues that this condition is necessary to obtain foveal fixation and thereby avoid artifactual summation. This condition obtained in the present experiment.

3. The possible effects of head and eye movements have been controlled in the present experiment by using a small stimulus patch presented to the area of the lowest gradient of retinal sensitivity, using a small bright fixation point in conjunction with artificial pupils, and using a mouth bite for rigid and reproducible head positioning.

4. Hewson (11), in her critique of Cook's experiment (3), argued that in order to avoid artifactual summation it is necessary to have both eyes fixated to insure foveal fixation. The results of Exp. 1 and 2 indicate that it is unimportant whether one or both eyes are fixated when the viewing condition is monocular. On the other hand, the results of Exp. 3 showed that, under conditions of binocular viewing, the P_R is significantly reduced when only one eye is fixated. For this latter finding at least two kinds of explanation can be proposed: (*a*) On the basis of Cook's argument (3) that a necessary condition for summation is equal sensitivity of the two eyes, the effect might be explained on the assumption that "wandering" of the unfixated eye results in stimulation of retinal points of differing sensitivity. This interpretation seems unlikely for two reasons. First, Crozier's results (4) on the retinal gradient of sensitivity indicate that large movements would be necessary. Results on the present apparatus (2) indicate that shifts as large as 20′ did not affect the P_R. Second, summation was found in the present experiment even when one of the monocular P_R's was zero. (*b*) A more likely explanation can be based on Crozier's argument (4) that a necessary condition for summation is the stimulation of corresponding points and the assumption that the "wandering" of the unfixated eye, if it occurs, results in stimulation of noncorresponding points. Support for this point of view is supplied by Shaad's results (16) in which she found that, at least, corresponding halves of the retinae must be stimulated for summation to occur.

5. Hecht's stimulus conditions for maximal retinal sensitivity (10) were used in the present experiment on the assumption that this would increase the sensitivity of the experiment for the effects to be investigated. The detection of summation cannot be attributed, without further research, to these optimal conditions, since an experiment by Pirenne (15) using almost identical stimulating conditions, with the exception of a longer duration of stimulation (.004 sec.), failed to find summation. This divergence of results might possibly be explained in

terms of Bartlett and Gagné's suggestion (1) that cumulative retinal changes produced by repetitive stimulation are a necessary condition for summation. In this respect the present experiment differed from Pirenne's and others in that it used a single luminance method of stimulus presentation with a fixed intertrial interval rather than one of the more conventional methods of threshold determination.

Since the conditions specified as necessary to avoid artifactual summation have been met, it seems plausible that summation has been obtained in the present experiment. On this assumption it is possible to evaluate the hypothesis proposed by Crozier (4, 5) and others that binocular stimulation of corresponding points of homogeneous areas of the retinae be considered as functionally equivalent to doubling the area of stimulation in a single eye. In order to evaluate the area-intensity function for the present data certain assumptions and approximations have to be made. Monocular left eye method of limits data for four sessions for each *S* were plotted on probability paper and straight lines fitted visually (2). The average slope over the four days was calculated for each *S*. From these slopes, on the assumption that the slope for binocular determinations would not differ from monocular determinations, the increment in log brightness necessary to produce the increment in percentage of "yeses" from the monocular to the binocular condition was calculated. The average for all *S*s and days was .18 log $\mu\mu$l with an *SD* of .03. The same value was obtained from constant method data. Ricco's law ($AI = k$) predicts an increment of .30 log $\mu\mu$l which is obviously too large. Piper's law ($\sqrt{A}I = k$) predicts an increment of .15 log $\mu\mu$l. On the assumption that .15 is the population value a *t* test of the difference between the obtained value and the population value failed to reject the null hypothesis ($t = 2.13$, $df = 8$, $p > .05$), but is very close to being significant. A possibly better estimate of the area-intensity relation can be obtained from Crozier's data (4, 5) on the relation between absolute threshold and area in which $\Delta I = CA^{-Z}$. He found $Z = .59$ for his peripheral data.[6] This would lead to a prediction of an increment of .178 log $\mu\mu$l. It obviously does not differ from the obtained value. The Wald-Graham (8, 19) area function predicts approximately the same value for the periphery as Crozier's function. If the relation and the value of *Z* found by Crozier can be accepted, the results of the present experiment lead to the conclusion that binocular summation acts in the same fashion as doubling the area in a single eye.

Intertrial association.—The hypothesis of no association was rejected in 9 out of the 12 conditions of the experiment. One significant difference in the amount of association was found in the intra-experiment comparisons and one in the inter-experiment comparisons. Because of the necessarily small sample size in the trial blocks, the power of the tests of significance is low, making tentative any conclusions about the lack of association under some conditions or the lack of a significant difference in the amount of association between conditions. Within this restriction, the lack of systematic effects within and between conditions leads to the conclusion that the effects of monocular versus binocular fixation or viewing are small, if present at all.

In order to interpret this conclusion, reference must be made to the empirical sampling study on the combination of independent nonrandom sequences. As would be expected, the result of combining two independent nonrandom sequences, two by two, was to produce a significant reduction in the amount of nonrandomness. Thus, if the sequential effect was generated purely peripherally, by some such factors as independent nonrandom

[6] Crozier's stimulus was a rectangular patch with a fixed angular height of 20.8°. The outer margin was 12.8° from the center of the fovea. Variations in area were achieved by moving the inner edge toward the fovea. The value of .178 was obtained by substituting in the equation, $\log \Delta I_0 = -Z \log A + C$. If this is the case, $\log \Delta I_0/\Delta I_0 = Z \log .5$. This value can only be an approximation, since Crozier found that Z was sensitive to the shape and retinal location of the stimulus patch.

eye movements, changes in fixation, or pupillary responses, it would be expected that the binocular condition would result in a reduction of the associative effect over that observed under the monocular condition. Since no such reduction of the amount of association appeared to occur under the binocular conditions, it may be concluded that the associative effect is, at least in part, nonperipheral in origin. This conclusion is supported by the fact that the manipulation of such peripheral factors as fixation and pupil size, and, by comparison with previous experiments, such factors as flash duration, stimulus patch size, and retinal position did not appear to effect the amount of association observed.

Summary

The effect of monocular versus binocular fixation and viewing conditions on probability of response (P_R) and the amount of intertrial association was investigated.

The binocular P_R was found to be greater than the monocular P_R under all conditions, and it was argued that this result was not an artifact of the conditions of observation.

Evidence was advanced for the hypothesis that stimulation of two eyes is functionally equivalent to doubling the area of stimulation in a single eye under the conditions of the present experiment.

On the assumption of the peripheral origin of the associative effect, the expected reduction of the intertrial associative effect was not found, and it was concluded that the effect is nonperipheral in origin.

References

1. Bartlett, N. R., & Gagné, R. M. On binocular summation at threshold. *J. exp. Psychol.*, 1939, **25**, 91–99.
2. Collier, G. H. An investigation of the Poisson frequency of seeing function and some possible variables governing the independence of successive responses for behavior at the visual threshold. Unpublished Ph.D. thesis, Indiana Univ., 1951.
3. Cook, T. W. Binocular and monocular relations in foveal dark adaptation. *Psychol. Monogr.*, 1934, **45**, No. 3 (Whole No. 202).
4. Crozier, W. J., & Holway, A. H. Theory and measurement of visual mechanisms. I. A visual discriminometer. II. Threshold stimulus intensity and retinal position. *J. gen. Physiol.*, 1938–39, **22**, 341–364.
5. Crozier, W. J., & Holway, A. H. Theory and measurement of visual mechanisms, III. ΔI as a function of area, intensity and wave length, for monocular and binocular stimulation. *J. gen. Physiol.*, 1939–40, **23**, 101–141.
6. Graham, C. H. An investigation of binocular summation: I. The fovea. *J. gen. Psychol.*, 1930, **3**, 494–510.
7. Graham, C. H. An investigation of binocular summation: II. The periphery. *J. gen. Psychol.*, 1931, **5**, 311–327.
8. Graham, C. H., Brown, R. H., & Mote, F. A. The relation of size of stimulus and intensity in the human eye: I. Intensity thresholds for white light. *J. exp. Psychol.*, 1939, **24**, 555–573.
9. Guilford, J. P. "Fluctuations of attention" with weak visual stimuli. *Amer. J. Psychol.*, 1927, **38**, 534–582.
10. Hecht, S., Shlaer, S., & Pirenne, M. H. Energy, quanta, and vision. *J. gen. Physiol.*, 1942, **25**, 819–840.
11. Hewson, C. B. An experimental study of "Binocular Stimulus Summation." Unpublished master's thesis, Univ. of Virginia, 1931.
12. Hoel, P. G. *Introduction to mathematical statistics.* New York: Wiley, 1947.
13. Ireland, F. H. A comparison of critical flicker frequencies under conditions of monocular and binocular stimulation. *J. exp. Psychol.*, 1950, **40**, 282–286.
14. Mueller, C. G. Quantum concepts in visual intensity discrimination. *Amer. J. Psychol.*, 1950, **63**, 92–100.
15. Pirenne, M. H. Binocular and uniocular threshold of vision. *Nature*, 1943, **152**, 698–699.
16. Shaad, D. Binocular summation in scotopic vision. *J. exp. Psychol.*, 1935, **18**, 391–413.
17. Verplanck, W. S., Collier, G. H., & Cotton, J. W. Nonindependence of successive responses in measurements of the visual threshold. *J. exp. Psychol.*, 1952, **44**, 273–282.
18. Verplanck, W. S., Cotton, J. W., & Collier, G. H. Previous training as a determinant of response dependency at the threshold. *J. exp. Psychol.*, 1953, **46**, 10–14.
19. Wald, G. Area and visual threshold. *J. gen. Physiol.*, 1937–38, **21**, 269–287.

(Received April 10, 1953)

Journal of Experimental Psychology
Vol. 47, No. 2, 1954

VOLUNTARY REST PAUSE BEHAVIOR IN A BLOCK-TURNING TASK AS A FUNCTION OF WRIST-CUFF WEIGHT

ROBERT F. BOLDT

Princeton University

AND DOUGLAS S. ELLIS

Armament Systems Training Research Laboratory, Human Resources Research Center

The present study is concerned with behavior in a situation where *S*, while working on a simple motor task, is free to take rests whenever desired. It investigates the effects of the effortfulness of the task on two aspects of such voluntary rest pause (VRP) behavior: (*a*) the number of responses prior to a VRP; (*b*) the mean rate of responding prior to a VRP.[1]

Results from two recent studies of motor-task effortfulness suggested the present study. In these two studies, although the effortful manipulation did not influence performance level, supplementary evidence suggests that the effortful manipulation did influence other aspects of *S*'s behavior. Eckstrand (**4**) varied the amount of spring tension against which the stylus had to be depressed in the pursuit rotor. In preliminary work he found that some *S*s could not work for the required length of time at the most effortful condition. Yet, in his final sample of relatively proficient *S*s, performance was not depressed at this most effortful condition. Ellis, Montgomery, and Underwood (**6**) varied the work-surface height of a block-turning task. They found that verbal expressions of hostility and requests to cease work were considerably more frequent at the most effortful (highest) height. However, as in the Eckstrand study, performance was not depressed at this most effortful condition.

These findings suggest that behavior exhibited in reaching a VRP might provide response measures sensitive to effortfulness. Such an expectation also appears to be supported by a portion of Hull's theoretical formulation. The relevant equation is (**7**, p. 286):

$$_{s}\bar{E}_{R} = {}_{s}E_{R} - (I_{R} + {}_{s}I_{R}) = {}_{s}H_{R} \times D - (I_{R} + {}_{s}I_{R})$$

This equation points up the possibility that, as an organism evokes a series of responses, a point may be reached where the cumulative inhibitory effects of such response evocation ($I_R + {}_sI_R$) offsets the excitatory effects ($_sE_R$). At such a point $_s\bar{E}_R$ (effective excitatory potential) would become so small that, if permitted, a VRP would ensue. This formulation can be applied to the present experiment, where independent groups of *S*s, each working at one of several levels of effortfulness, are under instructions to take rests whenever desired. Since Hull states that the amount of inhibition generated per response increases with effortfulness, one would expect,

[1] In addition to rate and number of responses, one could also consider time elapsed in reaching a VRP. However, determination of any two of these three variables automatically specifies the third, i.e., time = rate × number. Our choice of rate and number as the variables to be investigated reflects the predilection of inhibition theory (**7**, pp. 277–303) to deal with these variables rather than time.

assuming that the conditions are equivalent in ${}_{S}H_{R}$ and D, that a VRP would be reached more rapidly at the more effortful conditions.

However, the theory does not specify how this difference would be exhibited in behavior. As compared with less effortful conditions, *Ss* at the more effortful conditions could adjust in one of three ways: (*a*) maintain the same rate of responding, but perform fewer responses (constant rate adjustment); (*b*) perform the same total number of responses, but at a slower rate (constant number-of-responses adjustment); (*c*) decrease both the total number of responses and the rate of responding (compromise adjustment).

The present study, then, has two main purposes: (*a*) to determine if the behavior exhibited in reaching a VRP is sensitive to task effortfulness; (*b*) to determine whether *Ss* adopt a constant rate, a constant number-of-responses, or a compromise adjustment to the task effort variable.

Method

Subjects and experimental design.—Eighty *Ss* furnished 20 replications within four conditions of effortfulness. The *Ss* were male and female volunteers from sections of the introductory psychology course at Iowa State College, and were randomly assigned to effort conditions. Data were collected during the fall and winter quarters of the calendar year 1951.

Apparatus and task.—The task consisted of three ten-block rows from the block-turning portion of the Minnesota Rate of Manipulation Test. Four conditions of task effort were used. Lead weights, weighing either 0, .5, 1, or 3 lb. were inserted into cloth cuffs attached to *Ss'* wrists. To maximize effortfulness, *S* performed the task in a sitting position. With his elbows flush with the edge of the table, the edge of the task board furthest from *S* was placed under the tips of his outstretched fingers. The blocks were centered in front of *S*.

Procedure.—Each *S* was tested individually. Following preliminary instructions (*S* was told that the experiment was concerned with the influence of voluntary rests on productivity), *S* was given three 60-block practice trials. Trials 1 and 2 were given with no weight attached. During Trial 1, which was not timed, errors in work methods were corrected. Trial 2, which was timed, was inserted to provide a measure of *S*'s initial proficiency. Trial 3 was administered with the appropriate weight cuffs attached, and served to acquaint *S* with his experimental condition. The time interval between Trials 2 and 3 was approximately 120 sec.

Following the practice trials, *S* was given instructions designed to structure the experimental situation as follows: (*a*) the situation was to be of at least 10-min. duration; (*b*) *S* was to turn over as many blocks as possible during this 10-min. period; (*c*) *S* would do better if he would take a rest when he felt he needed it; (*d*) all rests would be of 60-sec. duration. Whenever *S* was working during this 10-min. period he was given knowledge of results every 30 sec. in the form of the number of blocks turned during the preceding 30 sec. No formal control was exerted over *S*'s rest period activity, although *E* tried to minimize discussion with *S* during these periods.

Results

Initial proficiency.—A simple analysis of variance of the initial proficiency scores (scores on practice Trial 2) showed that the four groups were equivalent, the F value being less than 1. Bartlett's chi-square test (**10**, p. 249) indicated this analysis to be appropriate.

Rate data.—A rate score for each *S* was computed by dividing the number of responses to a VRP by the time elapsed before a VRP. Rate means for the four groups are plotted in Fig. 1. Each point is based on 20 *Ss*. Period 1 refers to the work period before the first VRP; Period 2 refers to the work period leading up to the second VRP. No additional periods are included since some *Ss* took only two VRP's.

An analysis of variance of the rate data is presented in Table 1. The analysis is a "mixed" design of Type III in Lindquist's terminology (**9**, pp. 281–284). Application of Bartlett's chi-square test showed the cell variances to be homogeneous, suggesting

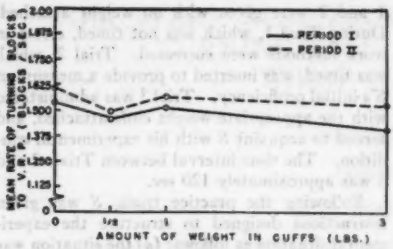


FIG. 1. Rate of turning blocks to a voluntary rest pause (VRP) as a function of effortfulness of response

that analysis of variance is appropriate. The weight mean square is barely significant at the 5% level of confidence. Period of practice also contributes significantly to the variation.

To discover which of the weight means is significantly different from the others, Duncan's multiple comparison test (3) was used. This test is designed to detect differences between all possible pairings of a group of means when no a priori choice of contrasts is available. The results of this test indicate that at the 5% level of confidence the 0-lb. group performed at a significantly faster rate than the .5-lb. group. No other differences were detected.

Number-of-responses data.—The mean number of responses before a VRP is plotted in Fig. 2. As before, each mean is based on 20 *S*s. Since the cell variances were heterogeneous,

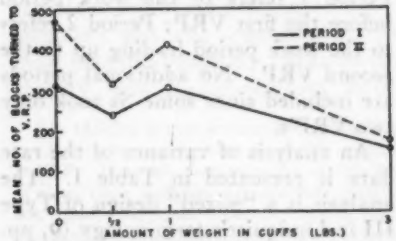


FIG. 2. Number of blocks turned to a voluntary rest pause (VRP) as a function of effortfulness of response

TABLE 1

ANALYSIS OF VARIANCE OF RATE DATA

Source	*df*	Mean Square	*F*	*p*
Weights (W)	3	.173	2.839*	.05
Stages of practice (SP)	1	.240	8.354**	.01
W × SP	3	.018	<1	
Subjects	76	.061		
Subjects × SP	76	.029		
Total	159			

* Tested against Subjects component.
** Tested against Subjects × SP component.

a log transformation was used on these data. These logs were analyzed in the same manner as the rate data. Since the variance of the transformed data was homogeneous, the analysis presented in Table 2 is appropriate. It can be seen that the difference between the weight conditions is significant at the 1% level of confidence. Duncan's test applied to these data shows that the 0-, .5-, and 1-lb. groups turned a significantly greater number of blocks prior to a VRP than the 3-lb. group. As was the case with the rate data, the period of practice effect is significant.[1]

DISCUSSION

The obtained results support the expectation that behavior exhibited in

[1] The data of the experiment can also be manipulated to obtain a reminiscence score for each *S*. Since each *S* took at least two VRP's, with a 60-sec. rest intervening between the first VRP and the start of work which eventually culminated in the second VRP, it was possible to estimate improvement over rest (reminiscence) by subtracting the time required to turn the last 30 blocks in Period 1 from the time required to turn the first 30 blocks in Period 2. An analysis of variance of these data yielded an *F* of 1.0. The result is consistent with current notions that reminiscence is an index of a temporary inhibitory process [temporary work decrement (1), reactive inhibition (8)] and that *S*s adjust to effortfulness so that inhibition is comparable at various levels of effortfulness (2, p. 99; 5). However, since a 10-min. rest is usually necessary to produce maximum reminiscence in motor tasks, results from a 60-sec. rest period do not provide very critical evidence for the hypothesis.

TABLE 2

ANALYSIS OF VARIANCE OF LOG NUMBER OF BLOCKS TURNED

Source	*df*	Mean Square	*F*	*p*
Weights (W)	3	.737	5.85*	.01
Stages of practice (SP)	1	.264	13.90**	.01
W × SP	3	.006	<1	
Subjects	76	.126		
Subjects × SP	76	.019		
Total	159			

* Tested against Subjects components.
** Tested against Subjects × SP component.

reaching a VRP does differentiate among effortful conditions. For the number-of-responses measure, the relationship seems clear: number of responses is inversely related to effortfulness, as shown by the significant differences between the most effortful condition and all other conditions.

For the rate data, the relationship with effortfulness is not so clear. One cannot be sure that there are, in fact, significant contributions to the total variance by the difference weights since the *F* value and the treatment contrasts were barely significant at the 5% level.

The *Ss*' verbal reports ("I was trying to beat my previous score") suggest that the change in both the rate and the number-of-responses measures between Periods 1 and 2 was primarily due to increased motivation in Period 2. There is, of course, the additional possibility of learning (e.g., *S* acquires more efficient response modes) being at least partially responsible for this effect.

Despite the inconclusiveness of the evidence relative to the change in the rate measure, the interpretation of the results of the experiment as a whole is clear. The behavior exhibited before a VRP is sensitive to effortfulness in the following matter: (*a*) the major adjustment is in the number of responses performed; (*b*) the amount of adjustment made in rate of responding, if any, is small.

Two important limitations on these results should be noted. First, the present study, like the Ellis, Montgomery, and Underwood study (6), used knowledge of results given in terms of the number of blocks turned every 30 sec., which is essentially a rate measure. Such rate knowledge of results may exert a stabilizing or restricting influence on *Ss*' rate-of-responding adjustment. Further study would be needed to determine if the role of number of responses as the primary adjustment to effortfulness represents a general principle of motor performance, or is specific to situations where rate knowledge of results is used.

A second limitation on these results derives from another, as yet unpublished, study which utilized the wrist cuffs. In that study conventional performance measures (rate of responding over a constant work period) were also sensitive to wrist-cuff weight. Thus the evidence is that VRP measures are sensitive to an effort variable to which conventional rate measures are also sensitive; i.e., this study may be considered a validation of VRP operations as an index of effortfulness. There have been, as yet, no investigations of VRP measures for effort variables like the work-surface height variable used by Ellis, Montgomery, and Underwood, which do not influence conventional rate measures. However, if VRP measures respond to work-surface height as they do to wrist-cuff weight, an explanation of why work-surface height failed to depress performance would be apparent. In that study *Ss* were not told the length of their work period, although it was essentially constant. The constant work period, which does not permit *S* to stop work after a certain number of responses, may deprive him of his primary adjustment (variation in number of responses performed) to effortfulness. Thus *S* would be forced either to allot more energy to the task or to adopt new work methods. The present results would seem to suggest that future research on effortfulness include VRP measures to check the possibility that *Ss*' methods of adjusting to a task are obscuring differences between effortful conditions.

The present results also suggest that VRP measures might be useful in studying recovery during rest as a function of

effortfulness. Reminiscence scores, when used in conjunction with effort variables like work-surface height which do not depress rate of performance, are difficult to interpret. It may be argued that they are based on response measures insensitive to effortfulness. However, a comparison of VRP measures before and after a single rest should provide recovery scores not subject to this difficulty.

Summary

Eighty *Ss*, 20 at each of four conditions of wrist-cuff weight, performed a block-turning task under instructions to take rests whenever desired. For each *S*, the number of responses made and the rate of responding exhibited in reaching a voluntary rest were determined. Analysis of the obtained data indicated that: (*a*) the number of responses made differentiated between the most effortful and all other effortful conditions; (*b*) the rate of responding did not consistently differentiate between effortful conditions.

These results were considered to be consistent with theoretical expectation. Their implications for experiments which study the effect of effortfulness on performance were considered.

References

1. Ammons, R. B. Acquisition of motor skill: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.*, 1947, **54**, 263–281.
2. Bilodeau, E. A. Decrements and recovery from decrements in a simple work task with variation in force requirements at different stages of practice. *J. exp. Psychol.*, 1952, **44**, 96–100.
3. Duncan, D. E. A significance test for differences between ranked treatments in an analysis of variance. *Virginia J. Sci.*, 1951, **2**, 171–189.
4. Eckstrand, G. A. Post-rest performance in motor learning as a function of amount of work performed in pre-rest practice. Unpublished Ph.D. dissertation, Northwestern Univ., 1949.
5. Ellis, D. S. Inhibition theory and the effort variable. *Psychol. Rev.*, 1953, **60**, 383–392.
6. Ellis, D. S., Montgomery, V., & Underwood, B. J. Reminiscence in a manipulative task as a function of work-surface height, prerest practice, and interpolated rest. *J. exp. Psychol.*, 1952, **44**, 420–427.
7. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
8. Kimble, G. A. An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.*, 1949, **39**, 15–23.
9. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Cambridge, Mass.: Houghton Mifflin, 1953.
10. Snedecor, G. W. *Statistical methods.* Ames, Iowa: Iowa State Coll. Press, 1946.

(Received June 1, 1953)

Journal of Experimental Psychology
Vol. 47, No. 2, 1954

SIMPLE CONCEPT LEARNING AS A FUNCTION OF INTRALIST GENERALIZATION [1]

MARIAN HOOPER BAUM

Yale University

Heidbreder has observed that when concepts of concrete objects, spatial forms, and numbers are learned in modified memory experiments, they tend to be attained in that order (6, 7, 8). She interprets this sequence as reflecting a gradient of "thing-character," or the extent to which an object has (*a*) constancy, (*b*) shape-boundedness, and (*c*) dynamic properties (8).

Gibson has deduced that the conventional concept-attainment experiment is a special case of paired-associate learning (3). The first list of stimuli and their respective responses constitute a series of paired associates. Subsequently presented lists consist of pairs in which each stimulus is "new" but is to be paired with a particular "old" response. One would expect that the generalization of each response from the first to the second list would reflect the extent to which the two stimuli (one from each list) are similar to or indistinguishable from each other.

But the effects of stimulus generalization in paired-associate learning are not limited to facilitation. Gibson has shown that retroactive inhibition increases with increasing interlist generalization (4); her finding that greater intralist generalization produces slower learning (5) has also been demonstrated by Underwood (11) and Underwood and Goad (12).

[1] This article reports a portion of a dissertation presented for the degree of Doctor of Philosophy at Yale University, 1951. The advice and encouragement of Professors Seymour B. Sarason, Carl I. Hovland, and John Dollard are gratefully acknowledged.

Each of four types of stimulus generalization may facilitate or interfere with concept learning: *Facilitation* may result from (*a*) interlist generalization among the various stimuli belonging to the same concept and (*b*) intralist generalization among various stimuli belonging to the same concept; this situation does not arise where each concept is represented only once in each list. *Interference* may result from (*a*) intralist generalization among various stimuli belonging to different concepts and (*b*) interlist generalization among various stimuli belonging to different concepts.

In the present analysis, both facilitating and interfering stimulus generalization in concept learning are assumed to be functions of stimulus discriminability, as this term is used by Gibson (3, 4, 5). To some readers the use of "dissimilarity," "differentiability," or "identifiability" may seem preferable to "discriminability" to designate a stimulus basis for what might be called generalizability. For purposes of the present argument, no assumption is made about the nature of the processes intervening between the discriminability of the stimulus and the generalizability of the response. Interference effects of generalization are measurable to the degree that they result in overt errors, i.e., nonsense names given as responses to stimuli other than those representing the concept for which they are "correct."

A hypothesis relating stimulus discriminability to order of concept attainments, then, may be stated as

follows: In any group of concepts learned concurrently, the ones which are composed of stimuli which are (*a*) least discriminable from each other and (*b*) most discriminable from other-concept stimuli in the task are the easiest to learn. Conversely, those concepts which are composed of stimuli which are most discriminable from each other and least discriminable from other-concept stimuli in the task are the most difficult to learn.

The present report describes a repetition of one of Heidbreder's classical, successive-list, concept-learning experiments. Here, in order to test one portion of the hypothesis proposed above, error data have been gathered. For each concept, the frequency of intrusion errors made by *S*s in learning responses to the first of 16 instances is used as an index of the intralist discriminability of all 16 stimuli comprising that concept. Our specific question is: To what extent is the order of attainment of each of nine concepts (each represented once in 16 different successive lists) related to the frequency of intrusion errors per stimulus in first-list learning?

Procedure

Twenty-seven randomly selected New Haven State Teachers College sophomores[2] were assigned at random to one of nine equal groups. All *S*s used the same concepts and stimuli in the same constant intralist orders, the same nonsense responses, and the same learning procedure. The intralist orders and the nonsense responses were taken from Heidbreder's Exp. B (8). The concepts in Heidbreder's 16 lists appear in various intralist orders. In the present study three different interlist orders, one for each group of *S*s, were used so that no single concept instance would appear in a given position for more than one-third of the *S*s. Here also concept-name pairings were varied for each group of *S*s in order to balance whatever advantage for learning, resulting from *S*s' pre-experimental experience, might accrue to any particular syllable to designate any particular concept.

The nine concepts, each represented by 16 different drawings, were: bird, face, hat (concrete objects); circle and two nonsense forms (spatial forms); and three, four, and six (numbers). The drawings used were copied from those of Heidbreder's Exp. B (8).[3] The nine response words, taken from the same experiment, were: fard, glif, joft, mank, palt, perg, quan, relk, and silm.

In individual session, each *S* first learned the nonsense names for a list of nine drawings, presented successively in constant order until *S* had reached the criterion of two consecutive perfect trials. Each drawing was exposed on a memory drum for 4.5 sec., including an anticipation period of 2 sec. The end of the anticipation period was signalled by a bell. After reaching criterion on the first list, *S* was given another list representing the same nine concepts; each drawing was different in form and list position from the first-list representation. For each drawing in the second and succeeding lists *S* learned one of the same set of nonsense names. The nonsense syllable "relk," for example, might designate the concept "face," being learned as a response to the drawing of a man's face in the first list, while in the second list "relk" was applied to a drawing of a woman's face. "Joft" for the concept "three" might be represented in the first list by a drawing of three umbrellas, while in the second list "joft" was learned as the response to a drawing of three curved lines. The *S*s were instructed that the task was memorizing and were never informed that concepts were to be attained. For each new list *E* waited for the end of the anticipation period before prompting *S*. Sixteen lists of nine drawings each were learned in this fashion.

The learning of each list was conducted according to Heidbreder's procedure: *E* immediately pronounced the name of the drawing as it appeared in the first list, and thereafter awaited the end of the anticipation period before prompting *S* when he did not respond. When *S* made an error he was corrected immediately. Each *S* was given three sets of instructions in the process of learning the 16 lists: once before the first list, again before the second list, and finally before the third list. These instructions were taken verbatim from Heidbreder's account of Exp. B (8) with two additions: (*a*) Preceding the first list *S* was told that a bell would signal the end of the anticipation period, and (*b*) preceding the second list *S* was given a list of the actual non-

[2] The author thanks Miss Ruth Rounds, sophomore adviser, for her invaluable assistance in procuring *S*s for the experiment.

[3] Professor Heidbreder generously made all her materials available for duplication.

sense names used in the experiment. These additions were made to the procedure and instructions in order to control rigidly the length of the anticipation period, and to standardize *S*'s pronunciation of the nonsense words as early in the task as was practical, respectively.

All verbal responses were recorded on a protocol sheet and on magnetic tape. Each *S*'s responses on first presentation of each new list, before prompting, were used to infer whether or not he had yet attained each concept. The relative difficulty of each concept was defined as the average number of different lists all *S*s learned to criterion before beginning to respond consistently and spontaneously to all succeeding drawings representing that concept with the correct nonsense name. This index is Heidbreder's "Concept Attained" measure.

The frequencies of invented syllables and of overt intrusion errors were analyzed according to the stimuli which evoked each, and to the ordinal number of list in which each occurred. The mean frequency of intrusion errors per drawing in the first list was defined as the intralist discriminability index of the stimuli comprising the concept it represented. Intrusions occurring in all trials on only the *first* list learned provide a readily available and clear-cut index of intralist discriminability; as soon as *S* sees a second and different list of drawings representing the same concepts, the effects of intra- and interlist generalization interference are not easily separated.

Results

In what order were the nine concepts attained? These data appear in the second column of Table 1. These attainments are generally consistent with Heidbreder's findings: the three object concepts are attained relatively early, the three spatial form concepts are attained next, and the three number concepts are attained relatively late. Using a *t* test which employs differences between each *S*'s scores on each two of the nine concepts, each number concept was more difficult ($p < .05$) than each form or object concept; the nonsense forms A and B were significantly more difficult than the bird and face concepts, but not significantly more difficult than the hat or circle concepts; and with the exception that the concept hat was significantly more difficult than the concept face, no differences between concepts of the same category were significant ($p > .05$).

To what extent are intralist generalization and speed of first-list learning related? Generalization interference effects were measured by first-list intrusion error frequencies per stimulus. The mean frequencies per first-list stimulus of correct responses, intrusion errors, and invented syllable

TABLE 1

Serial Positions, "Concept Attained" Scores, and Response Measures per Stimulus in the First List ($N = 27$)

Concept	Serial Position	"Concept Attained" Score		Correct Responses		Intrusions		Invented Syllables	
		Mean	*SD*	Mean	*SD*	Mean	*N**	Mean	*N**
Face	8, 6, 7	3.70	2.06	9.44	4.86	.18	4	1.63	17
Bird	3, 8, 1	4.11	2.44	9.07	3.61	.30	4	1.89	16
Hat	6, 3, 4	5.18	3.28	8.26	3.17	.33	7	1.89	19
Circle	5, 2, 3	4.56	3.00	8.92	3.41	.15	4	2.11	20
Form A	9, 5, 8	6.30	3.59	6.41	2.50	.67	11	1.81	15
Form B	2, 9, 6	6.26	3.42	7.78	3.43	.37	8	2.04	22
Three	7, 4, 2	8.78	4.50	6.89	4.35	.78	16	1.48	16
Four	1, 7, 9	10.44	4.63	7.07	3.37	.37	7	1.85	18
Six	4, 1, 5	9.59	5.03	6.52	3.09	.93	12	2.00	16

* These *N*'s represent the number of *S*s (of the total 27) making this kind of error in the presence of each stimulus.

TABLE 2

FREQUENCIES PER STIMULUS OF EACH INTRUSION ERROR IN THE FIRST LIST ($N = 27$)

Stimulus Concept	Response Concept									Total
	Face	Bird	Hat	Circle	Form A	Form B	Three	Four	Six	
Face			2		1				2	5
Bird			2			4	2			8
Hat	2	2			2		2		1	9
Circle					3	1				4
Form A	1		1			3	3	7	3	18
Form B		2			4		2		2	10
Three	1	3		2	5	5		2	3	21
Four				1		4	3		2	10
Six	3			3	3	5	6	5		25
Total	7	7	5	6	18	22	18	14	13	110

errors are also summarized in Table 1. (*SD*'s for neither type of error are presented; these distributions were so skewed that differences between means of these measures were not tested for significance.) These three classes of responses ("rights," intrusions, and invented syllables) comprise the total of first-list responses exclusive of omissions. One might expect intrusions to be related systematically to correct responses if for no reason other than the fact that any incorrect response must be an error. If this were so, intrusions and invented syllables should be (*a*) positively correlated and (*b*) about equally correlated (negatively) with correct responses. But Pearson's product-moment r (hereafter referred to simply as r) between intrusions and invented syllables per stimulus is $-.23$. And only the intrusion errors are significantly correlated with correct responses: r between intrusions and correct responses per stimulus is $-.86$ ($p < .01$), while r between invented syllables and correct responses is .10. The data imply that these errors reflecting the interference effects of intralist generalization alone are related systematically to the ease or difficulty with which responses are learned to these drawings comprising a single list.

Among what instances of the first list did *Ss* find discrimination most difficult? Table 2 shows the frequency of each intrusion error per stimulus. Inspection of these data reveals a heavy concentration of errors among the five most difficult concepts (two nonsense forms and three numbers). Intrusions in which one of these five serves as stimulus and another serves as response account for 67 of the total 110 errors made (in only 20 of 72 cells). The inference that these five concepts are responded to as being highly similar *to each other* appears to be justified.

To what extent is the order of concept attainment predictable from the interference effects of intralist generalization? Intrusions were found to be closely related to correct responses but were not significantly correlated with invented syllables within a single list. It remains to assess the relative

value of these two types of errors in predicting the eventual rate of attainment of each of the concepts. The relationship between order of attainment and degree of intralist generalization interference per concept, as measured by r between Concept Attained and first-list intrusion measures per concept is .69 ($p<.05$); r between Concept Attained and first-list invented syllables per concept is $-.12$.

Discussion

Present data show that when intralist generalization is measured by intrusion errors, the degree of this interference and the rate of learning per stimulus in the first list of this conceptual task are closely related. Invented syllables and intrusions per stimulus are not significantly correlated, and although invented syllables are more frequent than intrusions, intrusions are the more accurate predictor of correct responses. Finally, degree of intralist generalization per stimulus in the first list serves well to predict the order of eventual attainment of each of the concepts: the greater the generalization, the slower the eventual attainment. The data support the hypothesis that concepts represented by stimuli which are least discriminable from other concept stimuli in one list are the most difficult to learn, and that those represented by stimuli which are most discriminable from other concept stimuli in one list are the easiest to learn. For these materials, learned under conditions of the present study, where degree of generalization is measured by frequency of intrusion errors, that portion of our discriminability hypothesis which deals with intralist generalization interference is substantiated.

Confidence in accepting the hypothesis on the basis of these results hinges upon the justifiability of using errors as a measure of stimulus discriminability, or of whatever term one employs to describe the stimulus basis for nongeneralizability of response. In studies where discriminability has either been taken as a function of similarity along a physical dimension, such as height (9, 10), or has been determined pre-experimentally by judges' ratings (2, 4, 5), intrusion errors have been found to be highly correlated with discriminability. Newman's results (9) further suggest that errors, in contrast with number of concepts attained and number of correct responses, are the performance measure most sensitive to variations in similarity along a physical dimension. Taken together, these results seem to imply that intrusions generally increase with degrees of physical or judged perceptual similarity, but that such errors may under some conditions have no appreciable effect upon other performance measures. No a priori basis exists for ordering the Heidbreder drawings used here by either of these criteria of similarity. But the significant correlations obtained between errors and other performance measures suggest that the drawings vary systematically in discriminability over a wide range in a manner which corresponds closely to the speed with which responses to them are learned.

But is the order of attainment of these concepts simply a function of their discriminability? This criterion orders the stimuli in a sequence which also corresponds to that postulated by Heidbreder on the basis of their "thing-character." Since both criteria predict about equally well the attainment order of these concepts, using these drawings, a choice between hypotheses might ultimately be made on the basis of their generality. For general validity, the "thing-character" criterion should order attainments successfully despite variations in the nature of particular drawings used to represent these concepts, even despite variations among concepts chosen to represent corresponding degrees of "thing-character." Dattman and Israel's results (1) suggest that such generality of the "thing-character" hypothesis cannot be demonstrated. Whether or not the discriminability hypothesis has more ade-

quate generality is still to be tested experimentally. Present results support its interconcept intralist interference implications; an attempt to test the intraconcept interlist facilitation effects arising from the use of these stimuli is yet to be reported.

Summary

A stimulus-discriminability hypothesis is advanced to account for the ease or difficulty of attainment of each of nine concepts learned concurrently as paired associates. To test one implication of this hypothesis, that of the relationship between intralist interconcept generalization interference and order of attainment, one of Heidbreder's modified memory experiments was repeated.

Each of 27 *Ss* "memorized" a nonsense name for each drawing in 16 lists of nine drawings each. The nine nonsense names were applied consistently to different drawings representing each of nine concepts. The effects of order of presentation and concept-name pairings were balanced. Intralist discriminability was measured by the frequency of intrusion errors evoked by each stimulus in the first-list trial on each. The frequency per stimulus of invented syllable errors was tabulated.

The following are the major findings of this study:

1. The nine concepts were attained in approximately the same order as that found by Heidbreder: object concepts first, form concepts next, and number concepts last.

2. Intralist intrusions and invented syllable errors are not significantly correlated, but r between intrusions and correct responses per stimulus is $-.86$ ($p<.01$).

3. Of the two types of errors, only intrusions per stimulus predict eventual order of attainment per concept: r is $.69$ ($p<.05$).

These results are interpreted as supporting the interconcept intralist interference implications of a stimulus-discriminability hypothesis.

References

1. Dattman, P. E., & Israel, H. E. The order of dominance among conceptual capacities: an experimental test of Heidbreder's hypothesis. *J. Psychol.*, 1951, **31**, 147–160.
2. Gagné, R. M. The effect of sequence of presentation of similar items on the learning of paired associates. *J. exp. Psychol.*, 1950, **40**, 61–73.
3. Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.
4. Gibson, E. J. Retroactive inhibition as a function of degree of generalization between tasks. *J. exp. Psychol.*, 1941, **28**, 93–115.
5. Gibson, E. J. Intra-list generalization as a factor in verbal learning. *J. exp. Psychol.*, 1942, **30**, 185–200.
6. Heidbreder, E. The attainment of concepts: I. Terminology and methodology. *J. gen. Psychol.*, 1946, **35**, 173–189.
7. Heidbreder, E. The attainment of concepts: II. The problem. *J. gen. Psychol.*, 1946, **35**, 191–223.
8. Heidbreder, E. The attainment of concepts: III. The process. *J. Psychol.*, 1947, **24**, 93–138.
9. Newman, S. E. The effects of similarity and contiguity on the formation of concepts. Unpublished doctor's dissertation, Northwestern Univ., 1951.
10. Oseas, L., & Underwood, B. J. Studies of distributed practice: V. Learning and retention of concepts. *J. exp. Psychol.*, 1952, **43**, 143–148.
11. Underwood, B. J. Studies of distributed practice: II. Learning and retention of paired adjective lists with two levels of intralist similarity. *J. exp. Psychol.*, 1951, **42**, 153–161.
12. Underwood, B. J., & Goad, D. Studies of distributed practice: I. The influence of intra-list similarity in serial learning. *J. exp. Psychol.*, 1951, **42**, 125–133.

(Received May 4, 1953)

Journal of Experimental Psychology
Vol. 47, No. 2, 1954

INCREASED PRONOUNCING BEHAVIOR AS A FACTOR IN SERIAL LEARNING [1]

HELEN G. PRICE [2] AND DON LEWIS

State University of Iowa

The purpose of this study was to determine whether or not the serial learning of a list of three-consonant "syllables" of low association value is facilitated by attaching to them symbols which increase pronouncing behavior over that elicited by the consonants, but which give as little specific meaning to them as possible.

It has been well established that the degree of association value affects the readiness with which nonsense syllables are learned (4). Moreover, in a recent search in this laboratory for verbal materials which could be used in a special study of retroactive inhibition, it was found that most *Ss* were unsuccessful in learning pairs of three-consonant combinations of low association value by the paired-associates method. It appeared that those *Ss* who did learn parts of the lists were attaching pronounceable symbols to the consonant groups. A study of pronounceability as a possible factor in learning seemed indicated.

METHOD

Subjects and materials.—Six groups of undergraduate students in psychology classes, 15 per group, all volunteers, were used in testing the effect of added pronounceability upon the rate of learning a list of eight three-consonant combinations, the mean association value for which is 15.9%, according to Witmer's (6) ratings. The pronounceable symbols which were attached to the combinations were persons' names, none of which was known to refer to an actual person and which were therefore assumed to be of low meaningfulness. The "syllables" and their accompanying names appear in the third and fourth columns of Table 1.

Experimental design.—The experimental design is presented in Table 2. All groups first learned two lists of nonsense syllables, shown in Table 1, whose mean association values, according to Glaze (3), are 41% and 39.4%. This practice familiarized *Ss* with the serial anticipation method. The criterion was two perfect anticipations of the lists in any three consecutive trials.

The preliminary serial learning constituted the complete Day 1 treatment for Groups E and F. The other groups also received on Day 1 three exposures of each consonant combination of List III under the experimental conditions appropriate to each group.[3]

The experimental treatment for Groups A and B on Day 2 consisted of paired-associates training by the adjusted learning method (5). The *Ss* learned to associate a person's name with each consonant group to a criterion of three correct anticipations per combination. Following the paired-associates learning, Groups A and B learned List III serially. Group A reported their anticipations by using the names which they

TABLE 1

MATERIALS USED IN THE EXPERIMENT

Practice Lists		Experimental List III	Names Associated to List III Combinations
I	II		
QUB*	XIW*	ZXQ*	
NOL	JEP	BPJ	B. P. Jenkins
BUJ	MAF	GKZ	George K. Zook
GAF	VID	FJC	Fred J. Colby
CIH	BOJ	MHB	Martin H. Burke
DOQ	TEG	JGP	J. G. Pratt
TEB	FUW	HFK	H. F. Kelso
LUW	RIK	QCM	Quinton C. Manning
JIR	LAH	TLF	T. Lloyd French

* This combination was used only as a signal for the beginning of the list below it.

[1] This research was supported in part by the U. S. Air Force under Contract No. AF 33(038)-13214, monitored by the Perceptual and Motor Skills Research Laboratory, Lackland Air Force Base, San Antonio, Texas.

[2] Now at the Institute of Living, Hartford, Conn.

[3] In the course of exploratory work, it was found that the names were recalled more readily during the serial learning when the paired-associates learning was begun on Day 1.

TABLE 2

EXPERIMENTAL PROCEDURES

Group	Experimental Treatment	Criterion Task: Serial Learning List III
A	Paired-associates learning, List III combinations and names*	Response: names
B	As above	Response: initials
C	Paired presentations, List III combinations designated as initials and 8 colors per combination**	Response: initials
D	Paired presentations, List III combinations and 8 colors per combination**	Response: letter combinations
E	——	Combinations described as initials. Response: initials
F	——	Serial learning, List III. Response: letter combinations

* Criterion for true paired-associates learning: five correct anticipations of the response member of each combination.

** Each combination was presented a number of times equal to the mean number of exposures necessary for *Ss* of Groups A and B to reach criterion for learning for that combination in the true paired-associates task.

had associated to the combinations;[4] Group B responded with the consonants. Group B served as a control over the possible effect of knowledge that the consonants initials corresponded to specified names.

Groups C and D were also given experience with the List III combinations by the adjusted learning method. In the case of these groups, however, no associations could be formed, as each combination was paired with eight different patches of color (red, green, blue, yellow, gray, black, orange, and violet). Each combination was presented a number of times equal to the mean number of exposures necessary for *Ss* of Groups A and B to reach the criterion of learning for that combination in the true associated learning situation. Groups C and D could not anticipate the colors but were instructed to study the combinations carefully and to name the colors when they were exposed. Group C was urged strongly to think of the consonants as being persons' initials. Group D was instructed in terms of combinations of letters.

Groups C and D were included to determine whether or not familiarity with the consonant combinations would affect the ease of later serial learning. Group D was used as a control over the possibility that merely regarding the letters as initials would facilitate learning.

Groups E and F were the basic controls. They saw the consonant combinations in List III for the first time when they began the serial learning of them. Group E was urged to think of the consonant combinations as initials of persons; Group F was instructed in terms of combinations of letters. Thus, Group E was provided with a possible uniform basis for attaching pronounceable symbols to the "initials."

At the beginning of the serial learning of List III, all groups were given one presentation of the list during which they either called out the names as the combinations appeared (Group A) or called out the combinations themselves upon their appearance (Groups B, C, D, E, and F). This procedure gave the last named groups an opportunity to pronounce the responses required of them, a provision which seemed desirable inasmuch as Group A had been pronouncing its responses during the paired-associates learning.

The specific questions the experiment was designed to answer were:

1. Will the serial learning of consonant combinations of low association value be facilitated by prior training intended to attach to the combinations associations of negligible meaning but of undisputed pronounceability?

2. Will the serial learning show greater facilitation if *Ss* are required to respond vocally with the new associations during the serial learning than if they are required to respond with only the consonants?

3. Will familiarity with the combinations established during training prior to serial learning affect the serial learning?

4. Will knowledge that the consonant combinations may be classified as initials of persons' names provide a uniform basis for attaching pronounceable symbols to the combinations and consequently facilitate serial learning?

Apparatus.—A Missouri-type memory drum was used for the serial learning. The letter groups were presented in upper case pica typewriter type in vertical order; each list was arranged in a separate column on a wide sheet of

[4] Actually, any name was accepted as correct during the serial learning provided that the three initials involved corresponded to the consonant combination to be anticipated.

white paper mounted on the drum. The drum rotated at a constant speed, exposing the groups individually in the aperture of the apparatus at the rate of one every 3 sec., with an intertrial interval of 21 sec.

The experimental treatments were administered with the adjusted-learning device developed by Shephard and Lewis (5) which permits the presentation of each of any number of paired-associates as many times as desired and in any order. The stimulus member of a pair first appears alone in an aperture for 2 sec. A shutter is then raised to expose the response member for 2 sec. After a very brief delay, another stimulus member appears, starting the cycle again.

For all *Ss* trained on the adjusted learning device, the stimulus members consisted of the eight consonant groups comprising List III. The response members were either proper names whose initials corresponded to the paired consonant combinations (Groups A and B) or 3/4 in. squares of colored paper (Groups C and D). The consonant groups were typewritten full capitals in pica, the names in the conventional arrangement of upper and lower case pica. For Groups C and D, each consonant combination was paired with black, gray, blue, green, yellow, red, violet, and orange. The background for all pairs was white.

Results

For Lists II and III, the scores were the number of trials required to reach a criterion of two correct anticipations of the list in any three consecutive trials. The group means and *SD*'s appear in Table 3.

A simple analysis of variance was first applied to the measures for the various groups for List II. The obtained *F* was far from significant and the hypothesis that the scores were random samples from the same normally distributed population prior to experimental treatment can be safely retained.

TABLE 3

Number of Trials to Reach The Criterion

Group*	List II		List III	
	Mean	*SD*	Mean	*SD*
A	11.00	4.23	20.53	7.42
B	10.47	4.79	30.27	12.90
C	11.73	4.60	30.33	12.10
D	12.60	4.35	39.60	14.44
E	12.20	4.20	36.47	14.72
F	11.87	4.88	34.73	10.67

* $N = 15$ in each group.

TABLE 4

Differences between Means on List III†

Group	B	C	D	E	F
A	9.74*	9.80*	19.07**	15.94**	14.20**
B		.06	9.33*	6.20	4.46
C			9.27*	6.14	4.40
D				3.13	4.87
E					1.74

* Significant at the .05 level.
** Significant at the .01 level.
† The standard error of each difference, derived from the within-groups sum of squares, is 4.61.

A simple analysis of variance applied to the scores for the six groups for List III yielded an *F* of 4.18 which, for 5 and 84 *df*, is significant at the .01 level of confidence. Therefore, the *t* test for unrelated measures was used to evaluate the differences between pairs of means. Table 4 presents the differences, indicating those which are significant. For this study, the .05 level was deemed suitable for the designation of a difference as significant. It is apparent that the Group A mean differs significantly from all other means, and that the Group D mean differs significantly from the means for Groups B and C.

The data for List III were further examined in terms of the mean number of trials preceding the first correct anticipation of each consonant combination for each group of *Ss* during serial learning. The results are presented graphically in Fig. 1. The consonant "syllables" are indicated in the order of their appearance in List III along the abscissa. The ordinates depict the mean numbers of trials to the first correct anticipation of the eight separate "syllables." There is a separate curve for each group of *Ss*.

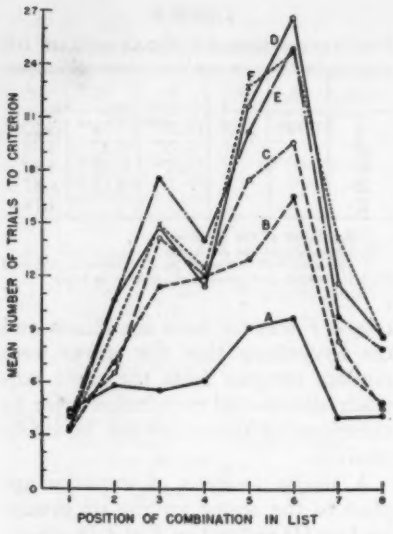


FIG. 1. Mean number of trials that preceded the first correct anticipation for each consonant combination. Letters refer to the six groups of *Ss*.

DISCUSSION

If the group means for number of trials to the criterion for the serial learning of List III are arranged in order of magnitude from low to high, the mean for Group A tops the list with the smallest number of trials. This group had established some familiarity with the combinations before their serial learning; was aware of their having significance as initials; had learned to associate specific names to the combinations; and, finally, this group pronounced the names during the serial learning of the combinations. Group F had none of these experiences.

A study of the serial learning performances of the other groups will assist in locating the source of Group A's superiority over Group F. Knowledge that the combinations were initials does not account for the superiority, inasmuch as Group E had this knowledge but did not perform more successfully than Group F.

Familiarity with the combinations developed before their serial learning does not explain the superiority of Group A. Groups C and D saw and studied each combination before the serial learning a number of times equal to the mean number of times it had been exposed to Groups A and B. Nevertheless, neither Group C nor Group D performed significantly better than Group F. Group D's significantly poorer performance with respect to those of Groups A, B, and C might have been due to the arousal of frustration. At the end of the experiment, many in this group expressed bewilderment and disbelief that nothing more was being required of them during the adjusted-learning procedure than that they study the combinations and call out the colors. Since they proceeded to the serial learning promptly after this treatment, frustration may have operated to interfere with the serial learning in a way described by Brown and Farber (1). Group C's slight tendency toward superiority over Groups E and F may be attributable to prior familiarity or to the combined effect of prior familiarity and the instructions to regard the combinations as initials, but no conclusion may be drawn which involves Group D data.

It is possible, if not probable, that attaching specific names to the combinations increased the latter's meaningfulness. That the superiority of Group A over the other groups is due to the combinations' increased meaningfulness is, however, doubtful. Groups A and B learned the same names, but Group B's performance is significantly poorer than Group A's and is not significantly better than Group F's. Correspondingly, the associative learning activity per se does not appear to have been facilitative, since Groups A and B both experienced it.

The only other single source of facilitation appears to be the extension of overt pronouncing behavior during the serial learning of the consonant groups. The *Ss* in Group A were the only ones permitted to pronounce the newly associated names during the course of serial learning, and their mean number of trials to the criterion was significantly lower than

that of Group B, as well as that of all other groups.

Since Group B was required to respond with letters only, it might be inferred that the process of rejecting the names and substituting their initials might have introduced some interference for this group. The slow (3-sec.) serial presentation rate, however, throws doubt on this inference.

If the decision to take .05 as the required level for significance is strictly adhered to, the discussion of the data in terms of number of trials to the criterion on List III must be concluded at this point. However, some attention should be given to the fact that the differences between Groups A and E and between Groups A and F are considerably greater than those between Group A and Groups B and C. These relationships suggest the possibility that the experiences of Groups B and C might have had some effects in the direction of facilitation, but of course no definite conclusions can be drawn.

The graphic presentation of the List III data in terms of the mean number of trials preceding the first correct anticipation of each combination for each group during serial learning also suggests that the experiences of Groups B and C, as well as those of Group A, were facilitative. It will be noted that the curves for Groups E and F, the basic control groups, show pronounced peaks to the right of the midpoint of the list, a characteristic common to the learning curves for difficult material (4). The curves for Groups A, B, and C approach the typical curve for the learning of easy material, their courses being flatter throughout most of the list and their peaks progressively reduced, from A through C.

These findings lend support to a hypothesis advanced by Gibson (2) which states that a large part of verbal learning consists of the development of responses which permit the discrimination of the units of the task. Groups A and B were provided with new associations to the consonant combinations which gave opportunity for making numerous already well-established responses which had previously served to differentiate among and identify verbal materials. The syllables and other elements of the names were familiarly observable, pronounceable, audible, and perhaps occasionally meaningful.

The obvious difference between the activities of Groups A and B was that Group A could make motor, auditory, etc. responses during the serial learning which had a long history of use in tasks of identifying and ordering, whereas Group B was forced to make again combinations of overt responses which were much less well learned. The slight tendency for Groups B and C to surpass Groups E and F may be due to the opportunity of the former groups to establish some identifying and organizing responses to the combinations during the adjusted-learning and pseudo-adjusted-learning experiences.

Summary

The study was aimed at determining the effects of increased pronouncing behavior on the learning of a list of eight three-consonant combinations by the serial anticipation method. After practice on the serial anticipation method, 90 *Ss* were assigned to six groups of 15 each.

Prior to serial learning of the consonant units, Groups A and B acquired pronounceable associations of slight meaningfulness (names of unknown persons) to the combinations by the paired-associates method. Adjusted learning procedures were employed. Groups C and D were presented the consonant combinations, each paired with eight colors, also by the paired-associates method. Group C was told that the combinations were initials of persons' names; Group D was not so informed. Groups E and F received no paired-associates training. Group E, but not Group F, was informed at the outset of serial learning that the combinations of consonants were initials.

All groups learned the list of eight consonant combinations by the serial anticipation method. Group A responded with (anticipated) the names previously learned. The other five groups spoke the consonants themselves, the instructions for Groups B, D, and F referring to the combinations as groups of letters, and instructions for Groups C and E referring to them as groups of initials.

Learning measures support the following conclusions: Increased pronouncing behavior was more facilitative when it was overt during serial learning (as it was for Group A), but some benefit accrued to serial learning when overt pronouncing was restricted to the paired-associates learning (as it was for Group B). The tendency toward facilitation noted in Group C's performance, perhaps due to familiarity with the consonant combinations, was not duplicated in the performance of Group D. Group D may have been adversely affected by the instructions received at the time of paired-associates learning. Mere knowledge (as for Group E) that the consonant combinations were persons' initials did not facilitate serial learning.

The findings were taken to indicate that the associated names provided a basis for making several kinds of well-learned responses which commonly assist in the discrimination and identification of verbal materials. In contrast, the consonant combinations by themselves were not favorable to the arousal of already established differentiating and identifying responses.

References

1. Brown, J. S., & Farber, I. E. Emotions conceptualized as intervening variables—with suggestions toward a theory of frustration. *Psychol. Bull.*, 1951, **48**, 465–495.
2. Gibson, E. J. Intra-list generalization as a factor in verbal learning. *J. exp. Psychol.*, 1942, **30**, 185–200.
3. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–267.
4. McGeoch, J. A., & Irion, A. L. *The psychology of human learning.* New York: Longmans, Green, 1952.
5. Shephard, A. H., & Lewis, D. A new device for presenting paired-associates in verbal-learning. *Amer. J. Psychol.*, 1950, **63**, 600–606.
6. Witmer, L. R. The association value of three-place consonant syllables. *J. genet. Psychol.*, 1935, **47**, 337–359.

(Received May 15, 1953)

Journal of Experimental Psychology
Vol. 47, No. 2, 1954

THE EFFECT OF PACING ON THE LEARNING OF A PSYCHOMOTOR RESPONSE

JACK A. ADAMS[1]

Perceptual and Motor Skills Research Laboratory, Human Resources Research Center

Paced and self-paced are two principal classifications of psychomotor tasks. In a paced task, the stimulus is presented for a specified time interval. At the end of this interval, whether *S* has made the correct response or not, the stimulus is withdrawn to reappear again or to be replaced by a new stimulus. In a self-paced task, no time limit is specified for stimulus presentation. Its presentation time is governed by the response speed of *S* and is only withdrawn when the correct response has been made.

Pacing is a differentiating performance variable only under certain circumstances. If the pacing interval, i.e., the duration of stimulus presentation, is sufficiently long, then a paced task will be equivalent to a self-paced task because all *S*s will have enough time to respond. Similarly, at high levels of skill where response latencies are short, *S*s will be able to complete a response sequence in a pacing interval that would have been inadequate at lower levels of learning. The pertinent question for motor learning would seem to concern the effects of certain pacing intervals at stages of practice where attainment of the task goal would be possible only part of the time. Under such circumstances pacing is obviously a *performance* variable since it operates to lower the number of completed responses, but it is not clear whether pacing also functions as a *learning* variable.

Mainly, there would seem to be two effects that pacing could have on the learning of a psychomotor response. First, because a paced task requires the response to be performed within a specified time interval, *S* will complete fewer response sequences required for achievement of the task goal, particularly at low levels of learning where response latencies are relatively long. If habit strength in a psychomotor task is a function of the number of times *S* achieves the task goal, then less "reinforcements" will be received under paced than under self-paced conditions and the level of habit strength in the task will be lower. Second, because a paced task emphasizes rapid response rate, *S* may acquire more efficient modes of responding to meet these task demands. This possibility is counter to the first in that it should lead to a higher level of habit strength for the task response under paced than under self-paced conditions. An appropriate test of level of learning would be to evaluate performance of paced and self-paced groups under common practice conditions. Obtained differences could then be taken to reflect learning differences and could be related to the degree of pacing in previous practice.

The present study was designed to evaluate the effects of pacing on the learning of a psychomotor response. The Complex Coordination Test, which is commonly employed as a self-paced task, was modified so that it could also be used as a paced task. Experimental groups, each practicing

[1] The experimental work for this study was performed as part of the United States Air Force Human Resources Research and Development Program. The opinions or conclusions contained in this report are those of the author. They are not to be construed as reflecting the views or endorsement of the Department of the Air Force.

with a different pacing interval, were shifted to self-paced practice in a series of test trials and compared with a control group having self-paced practice throughout. Performance on the test trials was examined for differences attributable to degree of pacing on prior trials.

Method

Apparatus.—The apparatus employed was a four-unit Complex Coordination Test, Model E. This test, as a self-paced apparatus, has been described elsewhere (2). To function as a paced task where the stimulus pattern would change after a specified interval, Hunter electronic timers were incorporated into the circuit to control the duration of stimulus presentation.

Subjects.—The *Ss* were 295 basic airman trainees drawn from the population available at Lackland Air Force Base, Texas. Because four units of the apparatus were used, subgroups of four *Ss* were assigned at random to the five main groups in the experiment.

Procedure.—All groups were administered 21 2-min. trials. So that comparability of groups might be determined, Trial 1 was self-paced for all groups. On Trials 2 through 11, a control group (C) continued to practice under self-paced conditions, and four experimental groups each had a different degree of pacing. The experimental groups and their pacing intervals were: I, 5 sec.; II, 7 sec.; III, 9 sec.; IV, 11 sec. An *S* in Group I, for example, had 5 sec. in which to match a setting. If he failed to match the setting during this interval, then the stepping switch automatically advanced and a new setting appeared. If, however, *S* matched the setting within the prescribed 5-sec. interval, then the stepping switch advanced as soon as the match had been made and he had 5 sec. within which to match the new setting. Trials 12 through 21 were test trials for evaluating level of learning, and all groups had self-paced practice.

The intertrial interval was 30 sec. in all instances but two. The interval between Trials 1 and 2 was lengthened to 1 min. to permit changing the apparatus to the paced condition for experimental groups and to give them brief instructions. These instructions were: "In the next series of trials, you will have —— sec. in which to match each setting. Keep trying to match as many settings as you can."

The interval between Trials 11 and 12 was increased to 3 min. to dissipate possible decremental effects attributable to work. No instructions were given to the experimental groups to indicate that practice would be self-paced on Trial 12.

The stepping switch was homed to No. 1 setting at the beginning of Trial 1 and Trial 12 so that all groups would begin practice on settings of equal difficulty at points where critical comparisons were to be made.

The *N*'s for each group were: C: 59, I: 58, II: 60, III: 59, IV: 59.

Results

Figure 1 shows the mean performance curves for the five groups. To determine comparability of groups prior to differential treatment, scores on Trial 1 were subjected to a simple analysis of variance to test the hypothesis that they were random samples from the same population. For 4 and 290 *df*, the obtained *F* ratio of .76 did not permit a rejection of the null hypothesis.

All groups are clearly differentiated on Trials 2 through 11. Group C, which had self-paced practice throughout, is clearly superior to the experimental groups which had paced training. Over-all mean levels of the experimental groups are ranked in accordance with the severity of pacing. Group I, which had a 5-sec. pacing interval, has the lowest mean level of performance and shows a trend which is essentially linear. Group II, with the 7-sec. interval, has a higher mean level than Group I and has a curve which is approximately linear with a slope somewhat greater than Group I's. Groups III and IV, with pacing intervals of 9 and 11 sec., respectively, have successively higher mean levels and their curves tend toward the curvilinear trend observed for Group C.

On Trials 12 through 21, all groups practiced under the self-paced condition. Figure 1 shows that the large between-groups differences present in Trials 2 through 11 have generally disappeared. A trend analysis was made to test the hypothesis that the mean

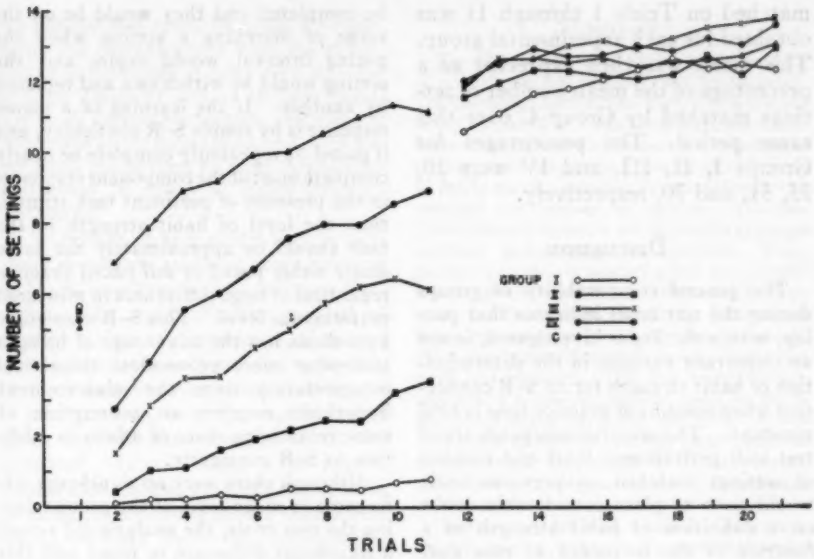


FIG. 1. Mean performance curves for all groups

levels and the slopes of the five curves over the test trials differ no more than could be attributed to random sampling. Table 1 gives the results of this analysis. To test the hypothesis that mean level on the ten test trials is the same for all groups, the pooled between-*Ss* mean square was used as the error term to evaluate the between-groups mean square. This gave an *F* ratio of 1.58 which, for 4 and 290 *df*, lacked significance at the 5% level. The hypothesis that the trends of the five performance curves are the same over the test trials involved testing the trials × groups mean square against the pooled trials × *Ss* mean square. An *F* ratio of 2.24 for 36 and 2,610 *df* was significant well beyond the 1% level of confidence. Examination of Fig. 1 suggests that this significant difference in trend is mainly attributable to Group I's curve having a somewhat steeper slope than the others in the early test trials. To evaluate this, the trials × groups mean square and the pooled trials × *Ss* mean square were computed with Group I deleted. The trials × groups mean square of 2.51 and the pooled trials × *Ss* mean square of 1.60 gave an *F* ratio of 1.57 which, for 27 and 2,097 *df*, was significant just beyond the 5% level and indicated that Group I was a principal contributor to trend differences.

TABLE 1

TREND ANALYSIS OF TRIALS 12 THROUGH 21

Source	*df*	Mean Square	*F*	*p*
Between groups	4	103.91	1.58	>.05
Between trials	9	64.50		
Trials × Groups	36	3.49	2.24	<.01
Pooled trials × *Ss*	2,610	1.56		
Pooled between *Ss*	290	65.64		
Total	2,949			

It is interesting to note how the pacing variable has affected number of settings matched prior to the test trials. The mean number of settings

matched on Trials 1 through 11 was obtained for each experimental group. This mean was then expressed as a percentage of the mean number of settings matched by Group C over this same period. The percentages for Groups I, II, III, and IV were 10, 25, 51, and 70, respectively.

Discussion

The general comparability of groups during the test trials indicates that pacing, within the limits investigated, is not an important variable in the determination of habit strength for an S–R connection when amount of practice time is held constant. The essential independence of test trial performance level and number of settings matched on previous trials would seem to place considerable doubt on a definition of habit strength as a function of the frequency of task goal achievement. Some notion of reinforcement could be invoked to explain the findings if reinforcement is defined in terms of knowledge of results of correct movements even though the movements are not quite sufficient to lead to a completed response. This interpretation can be reasonably applied to learning on the Complex Coordination Test since its stimulus panel provides knowledge of the extent and direction of each control movement. If an *S* under paced conditions fails to match a setting but has nevertheless matched both stick components and has initiated but not completed the correct rudder bar movement, his knowledge of the correct movements he has made could be interpreted as a reinforcing state of affairs. Thus, habit strength might be about the same as under self-paced practice even though few settings are actually matched.

The near equivalence of groups on the test trials is also consistent with an S–R contiguity learning hypothesis (1). Observation of *S*s practicing under paced conditions indicated that most of the component responses required in the total response sequence would frequently be completed and they would be on the verge of matching a setting when the pacing interval would expire and the setting would be withdrawn and replaced by another. If the learning of a motor response is by simple S–R contiguity, and if paced *S*s repeatedly complete or nearly complete most of the component responses in the presence of pertinent task stimuli, then the level of habit strength in the task should be approximately the same under either paced or self-paced practice regardless of large differences in measured performance level. This S–R contiguity hypothesis has the advantage of being a somewhat more economical theoretical interpretation since the reinforcement hypothesis requires an assumption of some reinforcing state of affairs in addition to S–R contiguity.

Although there were no significant differences in mean level of the groups during the test trials, the analysis did reveal a significant difference in trend and this difference was shown to be largely attributable to Group I, which received the most severe pacing. Even with Group I deleted, the hypothesis of some trend could not be unequivocally rejected. One possibility to account for trend differences is an uncontrolled variable that affected groups differentially. The arbitrary 3-min. rest interval between the final paced trial and the first self-paced trial could have been inadequate to dissipate work decrement effects accrued in prior trials. If work decrement is positively related to the severity of pacing, then Group I might be expected to show a different performance pattern over the test trials if the rest interval is too brief. Another possibility is that the pacing interval has a slight effect on habit strength. A brief pacing interval on the Complex Coordination Test should result in *S*s receiving little or no practice on the rudder control since it is a general tendency in the early stages of training to match the two stick components before the rudder. This lack of practice on an important component response could lower the habit strength level for the total response. An evaluation of these

two possibilities could be made by replicating the present study with the exception of introducing a long rest interval between Trials 11 and 12 to fully eliminate work decrement factors. Significant differences in performance trend over the test trials could then be more reasonably interpreted in terms of habit strength differences.

The results of this study bear importantly on certain applied human learning problems in industrial and military situations. The most effective method of learning a class of responses is, of course, of major concern in a training program. In many instances the problem of degree of pacing arises, i.e., the rapidity with which the stimulus material should be presented. In the absence of experimental data on this problem, it might be assumed on a priori grounds that severe pacing seriously limits the occurrence of completed response sequences with a consequent detrimental effect on learning. The present findings show, however, that pacing has a marked effect on performance but remarkably little effect on learning. This would suggest that a rapid presentation of stimulus material in a training situation could result in considerable learning despite the occurrence of few completed responses.

Summary

The effect of degrees of pacing on the learning of a psychomotor response was investigated. The Complex Coordination Test, which is commonly employed as a self-paced task, was modified so that it could also be used as a paced task. In four paced conditions *S* was given 5, 7, 9, or 11 sec. to match a stimulus pattern before it was removed and replaced by another. Ten test trials followed in which all experimental groups practiced under self-paced conditions. Comparison was with a control group having self-paced practice throughout. Principal findings were: (*a*) There was no significant difference in mean performance level over the test trials. (*b*) A significant difference in trend over the test trials was obtained. This was mainly attributable to a somewhat different trend for the 5-sec. group during the initial test trials. It was not clearly established whether these trend differences were attributable to habit strength differences or inadequate control of decremental work factors.

It was concluded that, within the limits investigated, pacing is not an important learning variable.

References

1. Guthrie, E. R. *The psychology of learning.* (Rev. Ed.) New York: Harper, 1952.
2. Melton, A. W. (Ed.) *Apparatus tests.* Report No. 4, Army Air Forces Aviation Psychology Program Research Reports. Washington: U. S. Govt. Print. Off., 1947.

(Received June 2, 1953)

Journal of Experimental Psychology
Vol. 47, No. 2, 1954

EXPERIMENTAL MANIPULATION OF VERBAL BEHAVIOR [1]

BERTRAM D. COHEN, HARRY I. KALISH, JOHN R. THURSTON, AND EDWIN COHEN

VA Hospital, Iowa City, Iowa [2]

The empirical basis of contemporary reinforcement theory consists largely of experimental data gathered with infrahuman *Ss*. Nevertheless, there is a strong desire among psychologists to extend these principles to human behavior. The recent texts of Dollard and Miller (1), Keller and Schoenfeld (3), McGeoch and Irion (4), as well as Skinner's treatment of verbal behavior (5) all include attempts at such extensions of theoretical coverage. It is obvious that the validity of these extensions can have important implications for social and clinical psychology. Of special interest in each of these treatments is the explicit role attributed to the principle of reinforcement in the modification of verbal patterns. Each of the writers mentioned assumes that a process similar to that typically observed in the instrumental conditioning of animal *Ss* operates to account for at least some proportion of human verbal behavior.

In an attempt to test this assumption, Greenspoon (2) introduced a conditioning procedure in which college students were instructed to voice nouns ad libitum. He found that the proportion of plural nouns elicited could be increased through reinforcement administered by having *E* say, "mm hmm" whenever such a response was made. The effect occurred without *Ss*' awareness of the contingency between their behavior and that of *E*. Another technique was developed by Taffel (6) in a study concerned, in part, with the effects of verbal reinforcement on the verbal responses of adult psychiatric patients. Taffel's procedure involved discrete stimulus presentations rather than the free-responding procedure used by Greenspoon. It required only that *S* be able to form a sentence when given a verb and a choice of one of six personal pronouns to be used in starting the sentence. Taffel found that the proportion of first person pronouns used by *S* could be reliably increased by verbal reinforcement without *S*'s awareness of the reinforcement operation. This technique is easily adapted to adult *Ss* who have varying degrees of verbal fluency and educational level.

The purpose of the experiments to be reported is the establishment of some basic relations between verbal reinforcement operations and verbal responses as elicited in a modified version of Taffel's procedure. The first experiment concerns the effect of verbal reinforcement on the acquisition of a pattern of pronoun preferences. The second experiment concerns the application of two extinction operations to the behavior of *Ss* following acquisition.

[1] A portion of this paper was read at the 1953 meeting of the Midwestern Psychological Association.

[2] Reviewed in the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published do not necessarily reflect the opinion or policy of the Veterans Administration. They are the results of the authors' own study.

EXPERIMENT I

Method

Apparatus.—The materials for Exp. I consisted of 80 3 × 5 white index cards. Eighty commonly used verbs were selected and a different verb was typed on each card. All the verbs were in the past tense. Six pronouns, I, We, He, They, She, and You, were placed below

each verb and their respective orders randomized for the series of 80 cards.

Subjects.—The *Ss* were 40 male ambulatory patients from the general medical population (excluding neurology and psychiatry) of the Veterans Administration Hospital in Iowa City. All *Ss* were under 55 years of age and capable of reading.

Procedure.—The 40 *Ss* were randomly assigned to control and experimental groups consisting of 20 *Ss*. Each *S* was informed that he was participating in a research project. The *S* was then shown a sample card and instructed to make up a sentence containing the verb indicated on the card and beginning with any one of the pronouns. Following this *S* continued to construct sentences until the series of 80 cards was completed. The *E* recorded each of the sentences verbatim. The cards were shuffled for each *S* to randomize the effect of order.

In the experimental group *E* said "Good" in a flat, unemotional tone at the end of any sentence which *S* began with I or We. In order to establish a base level for both groups without reinforcement, the reinforcement operation was not introduced until the beginning of the second series of 20 card presentations. The control group received no such reinforcement. In order to test *S*'s awareness of the relation between his responses and those of *E*, each *S* was asked at the completion of the series which pronoun was used most. Ultimately, all *Ss* were asked whether *E*'s behavior influenced their choice of pronoun.

Results

The data for Exp. I are illustrated in Fig. 1. The points plotted show the mean I and We responses for four successive blocks of 20 card presentations. The curves indicate a progressive increase in the frequency of the reinforced responses for the experimental group and no increase for the control group. A trend analysis yielded an *F* of 6.61 for the methods × trials interaction. This permitted rejection of the hypothesis of parallel curves beyond the .1% level for 3 and 114 *df*. The test for vertical displacement of the over-all means of these curves resulted in an *F* of 6.41 which is significant between the 1% and 5% levels for 1 and 38 *df*.

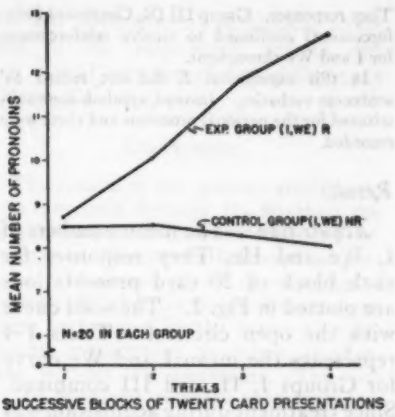


FIG. 1. Mean numbers of I and We pronoun for successive blocks of 20 card presentations

Questioning of *Ss* at the end of each series revealed no awareness of the contingency between their responses and those of *E*.

EXPERIMENT II

Method

Apparatus.—The materials for Exp. II were similar to those used in Exp. I. Since Exp. II was designed to investigate the extinction process, 80 additional verbs were introduced to provide four acquisition and four extinction trial blocks of 20 cards each. Thus, the total number of cards was increased to 160.

Subjects.—The 75 *Ss* in this investigation were also drawn from a general medical population and met the same selection requirements as *Ss* in Exp. I.

Procedure.—The *Ss* were randomly divided into three groups of 25 each. The procedure governing acquisition was similar to Exp. I for all groups with one modification which consisted of introducing reinforcement at the beginning of the first series of 20 cards. Five *Ss* were eliminated from the experiment because they showed marked position preferences, usually for the pronoun in the extreme left position.

The last series of 80 card presentations constituted the extinction phase of the experiment. During this phase, reinforcement of I and We responses was discontinued for Group I (SE, Straight Extinction) and Group II (RCR, Reinforcement of Competing Response). For Group II, however, reinforcement was shifted to He and

They responses. Group III (R, Continued Reinforcement) continued to receive reinforcement for I and We throughout.

In this experiment *E* did not record *Ss'* sentences verbatim. Instead, symbols were substituted for the personal pronouns and these were recorded.

Results

Acquisition.—The mean numbers of I, We and He, They responses for each block of 20 card presentations are plotted in Fig. 2. The solid curve with the open circles for Trials 1–4 represents the mean I and We curve for Groups I, II, and III combined. Since treatment during acquisition was the same for all three groups and no statistically reliable differences among the groups were found, the results were combined to form the single curve indicated. The anticipated progressive increase shown, similar to that found in Exp. I, is statistically reliable. The *F* of 15.23 for the main effect of trials is significant beyond the .1% level for 3 and 216 *df*.

The data plotted immediately below the I, We acquisition curve in Fig. 2 show the mean number of He and They responses for the same three groups combined. No statistically reliable differences among groups were found. The decrease in the frequency of He and They responses during acquisition of the reinforced response class was statistically evaluated. The *F* of 6.24 for the main effect of trials is significant beyond the .1% level for 3 and 216 *df*.

Extinction.—The values to the right of the line separating acquisition from extinction represent the mean number of I, We and He, They responses for each group when different treatments were introduced during extinction. The curve for Group III, the group which continued to receive reinforcement, shows no decrease in the frequency of the reinforced responses with an increase in the number of card presentations. It is also apparent that Group I, the group receiving no

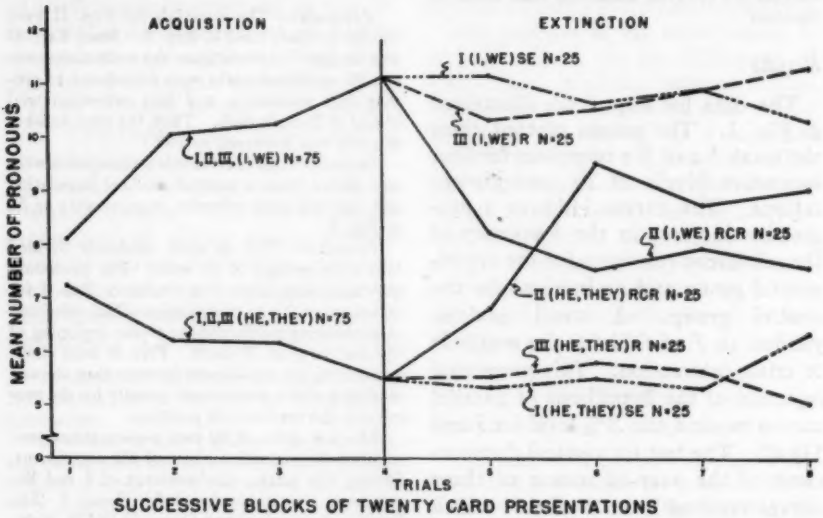


FIG. 2. Mean number of I, We and He, They pronoun classes for successive blocks of 20 card presentations during acquisition and extinction

reinforcement during this period, shows a marked resistance to extinction. This finding does not exclude the possibility that if the extinction trials had been extended beyond the 80-card series some decrease in the I and We response class might have occurred. Group II, in which the competing response class He, They was reinforced, shows a progressive decline in the frequency of the I, We responses. Trend analysis of the data suggests that the slopes of these three curves are not parallel. The *F* of 2.34 for the methods × trials interaction for Trials 4–8 is significant between the 1% and 5% levels for 8 and 288 *df*. Further analysis of the individual curves indicates that this interaction is due mainly to Group II. The test over Trials 4–8 for this group yields an *F* of 6.23 which is significant beyond the .1% level for 4 and 96 *df*. Similar *F*'s obtained for Groups I and III failed to exceed unity.

The data for the He, They response class shown in the lower half of Fig. 2 indicate no apparent change for Groups I and III, but a progressive increase in the reinforced He, They responses for Group II. The *F* of 4.62 for the methods × trials interaction is significant beyond the .1% level for 8 and 288 *df*. It is evident from an examination of these curves that the highly reliable methods × trials interaction is due to the increase in frequency of the He, They responses for Group II. The *F* of 10.12 for Trials 4–8 obtained for this group is significant beyond the .1% level for 4 and 96 *df*.

The use of four *E*s in the present investigation introduced the possibility of experimenter differences. In order to control this factor and to permit statistical tests, an approximately equal number of *S*s from each treatment group was randomly assigned to the *E*s. Subsequent statistical analysis of the acquisition results did not reveal any significant differences among *E*s.

Discussion

The results of the present study confirm previous findings (2, 6) that the principle of reinforcement does, in fact, operate to influence verbal response patterns. Moreover, direct questioning of *S*s indicated that the reinforcement operation, as applied in the present instrumental conditioning procedure, is effective without *S*'s awareness. The evidence cited also suggests that with this procedure the influence of the reinforcing stimulus is not limited to pronouns of self-reference (I, We), but can be effectively shifted to pronouns of other-reference (He, They).

One possible explanation for the failure to demonstrate extinction for Group I of Exp. II may involve the existence of continuing sources of reinforcement inherent in the procedure. Following each sentence, *E* continued to record *S*'s response and present the next stimulus. These events, in the absence of specific negative reinforcement, may carry sufficient reinforcement properties to maintain the existing response pattern. Such unspecific sources of stimulation seem to be implied by Skinner (5) in his hypothesis of "generalized reinforcement" for verbal responses (*tacts*) in everyday interpersonal interactions. The relatively marked change obtained for Group II, where the verbal reinforcer was shifted to another class of responses, indicates that the pattern does not become entirely fixed, but may be modified through appropriate shifts in the application of differential reinforcement.

The reinforcer in the present experiment ("Good") does not qualify, on operational grounds, as a secondary reinforcer since it was not experimentally paired with primary reinforcement. Within the conceptual framework of reinforcement theory, however, it may be considered as one of a large class of com-

plex secondary reinforcing stimuli including a great many verbal symbols of social approval (4). These stimuli probably acquired their reinforcing properties through association with primary rewards or through higher-order conditioning with other secondary rewards connoting security, affection, nurturance, etc., in the histories of the individual *Ss*.

Summary

Two experiments were performed to investigate the role of reinforcement in verbal behavior. In Exp. I two groups of 20 adult male *Ss* responded to a series of 80 cards containing a verb and six personal pronouns. The *Ss* formed sentences containing the verb and beginning with any one of the pronouns. For Group I all sentences beginning with I or We were reinforced by *E*'s saying "Good"; Group II received no reinforcement. The results, arranged in four trials of 20 cards each, indicated successive increments in the reinforced response for Group I and no change for Group II.

Experiment II concerned extinction following acquisition, and used 160 cards. The procedure governing acquisition was similar to Exp. I for three groups of 25 *Ss*. During extinction (the last 80 cards), reinforcement of I and We responses was discontinued for Groups I and II. For Group II, however, reinforcement was shifted to He and They responses. Group III continued to receive reinforcement for I and We. Analysis indicated no significant decrease in I and We responses for Groups I and III. Group II, however, showed marked extinction of I and We and a concomitant increase in the use of He and They. Questioning of *Ss* typically revealed no awareness of the contingency between their responses and those of *E*.

References

1. Dollard, J., & Miller, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
2. Greenspoon, J. The effect of verbal and non-verbal stimuli on the frequency of members of two verbal response classes. Unpublished doctor's thesis, Indiana Univ., 1951.
3. Keller, F. S., & Schoenfeld, W. N. *Principles of psychology*. New York: Appleton-Century-Crofts, 1950.
4. McGeoch, J. A., & Irion, A. L. *The psychology of human learning*. New York: Longmans, Green, 1952.
5. Skinner, B. F. William James' lectures on verbal behavior. Harvard Univ., 1948. Quoted in (3).
6. Taffel, C. Conditioning of verbal behavior in an institutionalized population and its relation to "anxiety level." Unpublished doctor's thesis, Indiana Univ., 1952.

(Received June 6, 1953)

Journal of Experimental Psychology
Vol. 47, No. 2, 1954

ANXIETY AND STRESS IN LEARNING: THE ROLE OF INTRASERIAL DUPLICATION

RICHARD S. LAZARUS,[1] JAMES DEESE, AND ROBERT HAMILTON

The Johns Hopkins University

A recent experiment by Deese, Lazarus, and Keenan (1) showed that there is an important interaction between score on a neuroticism questionnaire and the conditions of stress under which learning takes place. In that experiment the neuroticism inventory constructed by Winne (6) was administered to roughly 250 introductory and social psychology students. This inventory has a very high correlation with the Taylor anxiety questionnaire (4) and contains many of the same items. Approximately the upper and lower quartiles were randomly selected to serve as *Ss*. Low scores ranged from 0 to 2, high scores from 8 to 17. The resulting 90 *Ss* were divided into three groups of 30 each, each group containing 15 *Ss* scoring high on the inventory and 15 *Ss* scoring low. These three groups were randomly assigned to three experimental conditions. A different group of *Ss* was used for each of the experimental conditions so that the 30 *Ss* tested under one experimental condition were different *Ss* from the 30 *Ss* tested under the other conditions. The serial learning of nonsense syllables was compared under the three conditions in *Ss* characterized by high and low neuroticism scores.

Under control conditions, a small but nonsignificant difference was found between the learning of the high and low neuroticism groups. High neuroticism was associated with slightly more correct responses than low neuroticism. Under a second condition (avoidance), in which incorrect responses were followed by an electric shock, there was a very large difference between the high and low neuroticism groups. The high neuroticism group showed some facilitation in performance over the comparable control group. The low neuroticism group showed considerable impairment in performance as a consequence of the electric shock. In a third condition (nonavoidance), in which shock was administered randomly (that is, regardless of whether the *S*'s response was correct or not), no alteration of performance for the high neuroticism group was observed. However, the low neuroticism group suffered a marked decrement.

It was possible to interpret these results in terms of anxiety scores because of the close similarity between Winne's neuroticism inventory and Taylor's anxiety questionnaire. In fact, when the *Ss* in the Deese, Lazarus, and Keenan experiment were selected on the basis of their scores on Taylor's anxiety questionnaire, the relationships between anxiety score and performance remained the same as those found for neuroticism.

These results of this earlier experiment are consistent with those found by Taylor (4), Taylor and Spence (5), etc., on the conditioned eyelid response, but are in partial disagreement with the Montague (3) and Lucas (2) studies of rote serial learning. Montague found that when there was an opportunity for a high degree of intralist interference (that is, duplication of letters in the nonsense syllables), high anxiety was associated with

[1] Now at Clark University.

poorer performance. There is good reason to believe, therefore, that the nature of the task interacts with the relationship between anxiety and performance.

The Deese, Lazarus, and Keenan study (1) stands practically alone in that this relationship was studied under different conditions of experimentally produced stress. Moreover, one of the significant and unexpected findings of the study was that there existed a marked decrement under both stress conditions for the low neuroticism group. It is therefore important, in the context of the Deese, Lazarus, and Keenan study, to determine how these neuroticism or anxiety-performance interrelationships interact with the nature of the task under the experimental conditions of stress.

Procedure

The procedure in the present experiment was identical with that reported in Deese, Lazarus, and Keenan (1) with two exceptions. In the first place, the Taylor Anxiety Scale (4) was employed in the present study while Winne's Neuroticism Inventory (6) was used in the earlier experiment. It has been pointed out, however, that both inventories are highly correlated, and when the earlier data are reanalyzed for anxiety rather than neuroticism, the results are substantially the same. The scale was administered to a large introductory psychology class and approximately the upper and lower quartiles were selected. Eighty-five *Ss* were taken from this population and randomly distributed among the three experimental conditions so that each condition included equal numbers of both high and low anxiety *Ss*.

Second, the nonsense syllables used in the former study contained a minimal amount of duplication, while in the present case duplication was extreme. The list for the present experiment was composed of 12 consonant syllables using only five different consonants. This produced a task of great difficulty and marked intralist similarity.

Three different groups of *Ss* learned this list from a memory drum under three experimental conditions: one group was a control in which no shock or threat of any kind was employed; another was subjected to an avoidance condition in which *Ss* were shocked only for incorrect responses or failure to respond; and a third group received a random-shock condition in which shocks were administered equally often for correct and incorrect responses. In the random-shock condition the frequency of shocks was matched with their frequency under the avoidance condition. This arrangement was explained to the *Ss* before the experiment began.

In terms of the hypothesis that intralist duplication would result in a reversal of the anxiety-performance relationships, the present experiment may be compared with that of Deese, Lazarus, and Keenan.

Results

Table 1 shows the mean number of correct responses for high and low anxiety groups in all three conditions for both the present study and the earlier study cited above. It is immediately apparent that there is a large difference between the number of correct responses in the two experiments. The number of correct responses in the present study is about one-third

TABLE 1

Learning Scores (Correct Responses) in Present Experiment and in the Study by Deese, Lazarus, and Keenan

Condition	Deese, Lazarus, & Keenan Data				Present Data			
	Anxious		Nonanxious		Anxious		Nonanxious	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Control	31.40	15.52	27.87	16.76	10.00	5.74	10.36	4.67
Avoidance	37.07	20.88	19.53	8.94	12.21	5.86	13.20	5.59
Nonavoidance	27.53	19.67	20.47	14.04	10.64	5.54	10.87	7.22

that found in the earlier study using nonduplicated material. This indicates a great difference in difficulty for the two lists. Moreover, the number of correct responses in the present study is very similar to that found in the comparable trials in Montague's study.

The analysis of variance of the data for the present experiment shows that in contrast with the study by Deese, Lazarus, and Keenan (1), none of the differences is significant. The F for conditions is 1.37 with 2 *df* using the within-groups mean square as the error term. For anxiety and the interaction between anxiety and conditions the data are negligible with respect to any effects. However, the obtained differences with respect to anxiety are all in the predicted direction. The high anxiety groups are consistently though only slightly poorer than the low anxiety groups, and the best learning is found in the avoidance condition.

Discussion

Although the differences in the present experiment are not significant, the learning data are in accord with what is predicted from a combination of the experiments by Deese, Lazarus, and Keenan (1) and Montague (3). This judgment is based upon the subjective examination of the two experiments by inspection. There are three independent groups in the first experiment and three independent groups in the present study. While no statistical comparison has been made, the three groups in the first experiment are widely separated from those in the present study and the direction of the anxiety influence on learning is reversed. The nature of the task is an important determiner of the relationship between anxiety scores and performance.

It should be pointed out that Montague continued his *S*s beyond the point, in trials, at which the present study was terminated. Had the present study been carried as far as Montague's, it is likely that the differences would have been more marked. However, Montague did not employ a stress condition while the present experiment did. It was exceedingly difficult to get the *S*s to take as much shock as they did without protracting the painful experiment further. Actually, three *S*s refused altogether to submit to the experimental treatment after a few experiences with shock.

Elsewhere (1),[2] some of the present authors have held the view that the interpretation of this series of experiments might be more fruitfully considered in terms of other personality dimensions than anxiety. This interpretation was stimulated by the correlations between anxiety scores and scores on the hysteria and psychasthenia scales of the MMPI. Consequently, the present data were reanalyzed in terms of the hysteria-psychasthenia dimension. Psychasthenia scores on the MMPI for all *S*s were subtracted from hysteria scores (after overlap between hysteria and anxiety was eliminated), and the resulting distribution divided into two groups in terms of a median split. However, no significant change in the pattern of performance scores was found following this breakdown into psychasthenic and hysterical groups. The means and the analysis of variance for these data are presented in Table 2.

TABLE 2

Learning Scores (Correct Responses) for All Groups Based upon the Hysteria–Psychasthenia Distribution

Condition	High (Hysteria)		Low (Psychasthenia)	
	Mean	*SD*	Mean	*SD*
Control	10.25	4.51	10.08	5.98
Avoidance	13.75	5.42	12.79	6.05
Nonavoidance	11.00	7.42	12.09	5.11

[2] Deese, J., & Lazarus, R. S. The effects of psychological stress upon performance: a theoretical analysis and review of recent developments. Unpublished manuscript.

Summary

The present study is a repetition of an earlier study on the relationship between serial learning and personality variables under two conditions of stress produced by electric shock. In the earlier study marked differences were found between groups characterized as having high and low anxiety. The main basis of these differences was an impairment of performance by the low anxiety group under the stress conditions. In the earlier study, consonant nonsense syllables were used with a minimum of intraserial duplication. In the present study consonant nonsense syllables with a very high degree of such duplication were used. This difference in tasks produced a very large difference in results. No significant differences between the performance of the high and low anxiety groups were found. The obtained mean differences between the groups were actually in the opposite direction to that found in the earlier study. This is in accord with the results of Montague (3) and Lucas (2), and the hypothesis that task difficulty reverses the usual anxiety-avoidance learning relationships is supported.

References

1. Deese, J., Lazarus, R. S., & Keenan, J. Anxiety, anxiety reduction, and stress in learning. *J. exp. Psychol.*, 1953, **46**, 55–60.
2. Lucas, J. D. The interactive effects of anxiety, failure, and intraserial duplication. *Amer. J. Psychol.*, 1952, **65**, 59–66.
3. Montague, E. K. The role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, **45**, 91–96.
4. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
5. Taylor, J. A., & Spence, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, **44**, 61–64.
6. Winne, J. F. A scale of neuroticism: an adaptation of the Minnesota Multiphasic Personality Inventory. *J. clin. Psychol.*, 1951, **7**, 117–122.

(Received May 25, 1953)

Journal of Experimental Psychology
Vol. 47, No. 2, 1954

INTELLECTUAL PERFORMANCE AS A FUNCTION OF TWO TYPES OF PSYCHOLOGICAL STRESS[1]

SONIA F. OSLER[2]

The Johns Hopkins University

In a recent review on the effects of psychological stress on performance, Lazarus, Deese, and Osler concluded that "research in . . . [this area] must start practically afresh" (6, p. 315). This rather startling statement was made in the interest of a systematic approach to the study of the variables involved. The need for such a systematic approach was emphasized by the discrepant results so far reported by investigators in this area—results which varied all the way from demonstrating improvement to impairment in performance—and by the absence of criteria for the prediction of individual performance under stress.

The present investigation represents an initial attempt at a systematic study of the effects upon performance of different stimulating situations used to produce a state of stress. A variety of situations have been utilized in past research, among them failure (1, 3, 12), excessive pacing (1, 8), distraction (10), and threats of electric shock (2, 17). However, little attention has been paid to the possibility that the variations in the stimulating conditions may have been responsible for the differences in results obtained. Stopol (13) reports an experiment in which he attempted to manipulate this variable. He exposed his subjects to two different forms of stress and observed no differences in group effects. In most stress experiments the emphasis has been placed on the generally affective aspects of the situation rather than the operations used to produce it. It has been implicitly assumed that the specific conditions differed little, if at all, in their effect upon performance.

It is this last point that has been the primary concern of the present study. Here the effects of two types of stress situations upon intellectual performance have been investigated. One of these consisted of the thwarting of the motivation to succeed in the experimental task. This was represented by the already familiar failure-stress situation. The second type of stress situation involved the thwarting of a motive not directly related to the performance which was measured. This was represented by a disturbing situation unrelated to success or failure on the experimental task.

The hypothesis proposed is that the effect on performance of each of these two situations is not the same.

An attempt will be made to answer two additional questions: first, whether there is any relationship between the subjective reactions to the stress situation and performance; and, second, whether there is any relationship between *S*s' usual (i.e., prestress) level of manifest anxiety and the effect of stress upon their performance.

There are very few studies pertinent to the first question (14). So far as anxiety is concerned, recent research by Taylor and others (2, 15, 16) indicates that high manifest anxiety is associated with more rapid avoidance

[1] A dissertation submitted to the Faculty of Philosophy of The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

[2] The author is grateful to Dr. R. S. Lazarus for his guidance and criticism of this work, and to Dr. C. W. Eriksen for his helpful suggestions.

conditioning and simple learning than low anxiety. Investigations by Montague (11) and Lucas (9) indicate that when the tasks become more difficult anxiety is associated with lowered performance. A study by Lazarus, Deese, and Hamilton (5) produced no significant differences in performance on tasks of varying difficulty. Lazarus and Deese,[3] using military personnel, report results opposed to Taylor's. In the present experiment, which utilizes a difficult task, the expectation is that high anxiety should be associated with poorer performance.

Procedure

This experiment was conducted in three Baltimore high schools[4] and involved 15 intact eleventh-grade classes. Altogether 331 students participated, but because of lateness for one or both of the experimental sessions, absences of partners in a matched group, or unscorability of tests, the final analysis is based upon 305 *Ss*.

The experiment took place during two school periods of 50 min. each. It was conducted by the teachers who were coached by *Es*[5] in the exact procedure to be followed. This was done in the interest of maintaining a realistic situation in which *Ss* were more likely to believe the stressful information.

The task was an arithmetic test consisting of 30 problems in long division. Each problem called for the division of a six-digit number by a three-digit divisor to yield a three-digit quotient with a two-digit remainder. Several precautions were taken to keep the problems equally difficult. In no case was there a digit "1" and in only one case was there a repetition of a digit in the quotient, since this would render the problem easier. The problems had been previously equated as to difficulty on a group of college students and on a group of tenth-grade high school students. Two equivalent forms of this test were constructed.

At the first experimental session *Ss* were given the following instructions:

"Today we will do something different. The Department of Education is conducting an evaluation survey in all Baltimore high schools. You will be given a test today which is part of the survey."

No further information was given. Questions as to whether the grade obtained on this test would appear on *S*'s record were either evaded or answered: "I don't know." Actually, very few questions were asked by *Ss* at this session.

Twenty-five minutes were allowed for the test. During the remainder of the 50-min. period *Ss* filled out a questionnaire which consisted of 50 items from Taylor's anxiety inventory (15).

Before the next experimental session *Ss* from 12 of the 15 classroom groups were divided into four matched groups on the basis of the number of problems completed on the first test. The three classes not pooled, one in each of the high schools, constituted one of the control groups.

In order to avoid biasing effects that might arise from differences in schools or particular classrooms, the experimental groups were so formed that each class contributed approximately one-fourth of its students to each of the four experimental groups.

One week following the first session *Ss* in the four matched groups were given the following instructions:

"Today we will continue with the Department of Education test. Attached to the paper you will receive today you will find your grade on last week's test, so that you may know where you stand so far. These tests are considered very important, and only those students who do very well may be recommended for college. Do your very best."

The *Ss* in the five groups were given the following treatment. Groups I, II, III, and IV

[3] Unpublished data.

[4] The author wishes to acknowledge with gratitude the active cooperation of the Department of Education of the City of Baltimore, particularly Dr. J. C. Taylor, Assistant Superintendent of Schools; Dr. A. Lichtenstein, Director, Division of Special Services to Pupils; Dr. C. H. Katencamp, Principal, and Mr. H. T. Yost, Vice-Principal, Baltimore City College; Miss A. M. Zouck, Principal, and Miss A. Miller, Vice-Principal, Eastern High School; Mr. W. E. Dunn, Principal, and Mr. C. Whitney, Counselor, Forest Park High School; and the 15 classroom teachers who were in direct charge of the two experimental sessions.

[5] The advisability of conducting the experiment simultaneously at the three schools, in order to prevent exchange of information, necessitated that three different *Es*, one at each school, supervise the conduct of the experiment. The author is indebted to Dr. Richard S. Lazarus and Mr. Virgil R. Carlson, who each supervised at one school and instructed the teachers in the conduct of the experiment, and to Mr. Philip Teitelbaum who assisted the author in the largest of the three schools in which the experiment was conducted.

received all communications by means of written notices attached to their respective test papers. All notes were written on standard school forms that would normally be used for a similar purpose. They were filled out and signed in the manner customarily followed in each school. The note referring to the grade on the arithmetic test was especially designed for the purpose. It was mimeographed, bore the Department of Education heading, and was rubber-stamped with the information relating to the test performance. This produced quite an "official" looking note. Below is a description of the treatments given the five experimental groups.

Group I: These *Ss* received a note reading as follows: "You failed Evaluation Test IIA.[6] You were in the lowest 10% of all the students so far tested. This is a very poor grade. Do your best in this test." This group shall be referred to as the failure group.

Group II: These *Ss* received the same failure notes as *Ss* in Group I, and in addition they received a note reading as follows:[7] "Please report to the office at the end of this period. We have received a serious complaint about you. Your family has been notified." The note bore the signature of the individual who normally handled disciplinary questions in each school, usually the vice-principal. This group shall be referred to as the failure-fear group.

Group III: These *Ss* received the same "disciplinary" note as Group II, and in addition they received the following note: "You passed Evaluation Test IIA. You did very well last time as compared with all the students so far tested in Baltimore. Do your best again in this test." This group shall be referred to as the success-fear group.

Group IV: These *Ss* received the same passing note as *Ss* in Group III, and no other information. This group shall be referred to as the success-control group.

Group V: These *Ss* (the three intact classes) were given the following instructions orally by their respective teachers: "Last week you took a test in arithmetic. This was not a real test but simply part of an experiment in how quickly and accurately people can work under different conditions. Today I will ask you to do similar problems again. The results you obtain will count neither for nor against you. But you are asked to work as quickly and accurately as you can, as we are interested in obtaining a measure of best production from you." This group shall be referred to as the control group.

[6] This was the code number assigned to the first test.

[7] The wording differed in each school to conform to the form actually in use.

In the 12 classrooms in which *Ss* were not treated uniformly, every effort was made through a special seating arrangement to keep *Ss* from knowing that anyone else in the classroom had received a note regarding the disciplinary matter or to learn anyone's standing in the previous week's test.

During the course of the second experimental period the teachers kept careful notes of any overt signs of distress. At the end of 25 min. the test was terminated; *Ss* in Groups I, II, III, and IV were then informed that they had just participated in a psychological experiment. They were requested to report to the auditorium where the nature of the experiment would be explained and an opportunity given for questions and discussion. The *Ss* in Group V were also asked to report to the auditorium so that they, too, had the opportunity to learn more about the experiment.

At this meeting lively discussion ensued at all three schools. Following this *Ss* were requested to fill out a questionnaire which dealt with their subjective reactions to the experiment. It asked for information as to the extent to which they believed the information attached to their test papers, whether they were concerned over it, and how they thought that concern affected their performance.

Method of scoring the arithmetic test.—Each example required six steps for correct completion, and each of these steps was given one point credit, so that an entire example, if completed correctly, was given six points. If, because of one error in calculation, subsequent steps in the problem differed from the key answer but nevertheless represented operations performed correctly, credit was given for such work. In other words, an error was penalized only once. However, if the digit "1" appeared in the quotient because of a previous error, or if any other digit appeared twice in the quotient (except for the one problem previously noted), the steps in the example calling for multiplication by "1" or by a repeat digit received no credit.

Results

Realistic nature of the stress situation.—A crucial aspect of any stress experiment is the creation of a realistic situation. It was important, therefore, to examine the available information regarding this aspect of the present study. The teachers' observations, the meetings conducted in each school, and the answers to the written questionnaire revealed that

TABLE 1

MEANS AND *SD*'S OF SCORES OBTAINED ON THE TWO TESTS
($N = 61$ in each group)

Group	Test I		Test II	
	Mean	*SD*	Mean	*SD*
I. Failure	71.3	28.1	79.2	27.3
II. Failure-fear	70.9	28.9	77.5	29.6
III. Success-fear	72.3	26.2	81.9	27.7
IV. Success-control	71.0	30.4	84.1	31.8
V. Control	63.8	22.0	77.5	23.4

most *Ss* were deeply involved in the situation. In some instances they voiced resentment at having been subjected to such a threatening experience. It is of special interest to note that while 87% of the stressed *Ss* indicated that they believed the notes, 92% stated that they worried about them. Some *Ss* who stated that they did not believe the notes, suspecting that this was an experiment, added that they could not help worrying about them. The proportion of *Ss* in each of the four experimental groups who believed the notes was as follows: failure, 89%; failure-fear, 92%; success-fear, 82%; and success, 84%.

Reliability.[8]—The correlation coefficient between scores obtained in the two tests for all *Ss* was .89. Within groups they varied from .87 to .93. The means and *SD*'s of the five groups in each of the two tests are shown in Table 1.

The results of the analysis of covariance of the four experimental groups can be seen in Table 2. These show

[8] It will be recalled that Groups I, II, III, and IV were matched on the number of examples completed in the first test rather than the actual scores achieved. The correlation coefficient between the number of examples completed on the first test and the score received on that test is .98. The corresponding coefficient for the second test is .99. From this it can be deduced that accuracy of performance remained the same for the two tests.

TABLE 2

ANALYSIS OF COVARIANCE—FOUR EXPERIMENTAL GROUPS†

Source	*df*	SS of Errors of Estimation	*F*
Combined‡ error	180	28,823.0	
Failure plus error	181	29,879.1	
Adjusted SS-failure	1	1,056.1	6.60*
Fear plus error	181	28,900.5	
Adjusted SS-fear	1	77.5	
Levels plus error	240	39,081.6	
Adjusted SS-levels	60	10,258.6	1.07

† In this analysis adjustment was made for differences in initial arithmetic scores.
‡ All three first-order interactions were not significant. They were therefore combined with the error term and the degrees of freedom were added.
* Significant at the 5% level of confidence.

that failure as an experimental variable depressed performance, the difference being significant at the 5% level of confidence. The other experimental stress, fear, did not affect performance significantly.

In order to include the data from the control group an analysis of covariance was performed with the breakdown now consisting of groups and levels of initial ability. Table 3 shows that the variance due to groups is significant at the 5% level.

Group-by-group comparisons were made in accordance with the method suggested by Goulden (4, p. 157). On the basis of the adjusted means as shown in Table 4 the difference between the failure and success-control groups is significant at the 5% level;

TABLE 3

ANALYSIS OF COVARIANCE OF THE ARITHMETIC SCORES OF THE FIVE GROUPS

Source	*df*	SS of Errors of Estimation	*F*
Error (residual within groups)	239	35,722.4	
Between groups plus error	243	37,422.1	
Adjusted SS between groups	4	1,699.7	2.84*

* Significant at the 5% level of confidence.

the differences between the failure-fear and the success-control groups and between the failure-fear and the control group are significant at the 1% and 5% levels, respectively. No other differences are significant.

Subjective reports.—In order to determine whether subjective reports were related to arithmetic performance the median test was used. The difference in performance between those who said that they believed the information contained in the notes and those who said that they did not believe it was not significant. The same was true about the difference between those who said they worried and those who said that they did not worry. It should be noted that in each case 82% or more of the *Ss* fell into one subjective category.

Anxiety and performance.—In order to determine whether *Ss'* anxiety level was related to performance under stress, correlation coefficients between the anxiety score and the difference in performance on the two tests were calculated for each group. Table 5 lists these correlations for the five groups and for all *Ss* combined. None of these correlations is significantly different from zero, though it is interesting that the three stressed groups show negative correlations and the two control groups show positive correlations. The probability of such a finding occurring by chance is less than .05. The groups did not differ in average initial anxiety level. The correlations between anxiety score and scores on Test I and Test II were not significantly different from zero.

TABLE 4

ACTUAL AND ADJUSTED FINAL MEAN ARITHMETIC SCORES, ALLOWING FOR THE DIFFERENCES IN INITIAL ARITHMETIC SCORES

Group	Obtained Mean	Adjusted Mean*
I. Failure	79.2	78.1
II. Failure-fear	77.5	76.7
III. Success-fear	81.9	80.1
IV. Success-control	84.1	83.2
V. Control	77.5	82.0

* A difference of 4.5 is significant at the 5% level and a difference of 5.9 is significant at the 1% level of confidence. Standard error computed in accordance with Goulden (4, p. 158).

TABLE 5

CORRELATION COEFFICIENTS BETWEEN ANXIETY SCORE AND DIFFERENCE IN PERFORMANCE IN THE TWO ARITHMETIC TESTS

Group	r
I. Failure	−.107
II. Failure-fear	−.150
III. Success-fear	−.195
IV. Success-control	+.157
V. Control	+.051
All groups combined	.000

DISCUSSION

It is apparent that the effect of psychological stress upon intellectual performance is not only a function of the type of stimulus situation used to create a stressful situation, but also of the type of control situation used for comparison. Thus, failure as a stimulus depressed performance significantly, but fear did not. Further, from the individual comparisons we see that the failure group differs significantly in performance from the success-control group, but does not differ from the control group.

Perhaps it should be pointed out that while the original intention had been to compare the effects of fear with the effects of failure, unfortunately the fear groups were contaminated with the latter. It might have been more desirable had one fear group received a passing note that did not imply exceptional success or failure, so as to have provided an opportunity to observe the effect of fear relatively uncontaminated by other factors. However, such was not the case.

What may be responsible for the differences in effects produced by the two types of stress? At best we can only conjecture on this point. Two possible

explanations present themselves: one in terms of the differences in the intensity of stress, and the other in terms of the differences in the relation of the stress to the experimental task.

On the supposition that the two types of stress merely differed in intensity, the findings of this experiment are in line with Glixman's (3) on the recall of completed and incompleted tasks under varying degrees of stress. However, the subjective reports of *S*s would call to question a conclusion that fear was a less intense stress than failure.

We are inclined to account for the results in terms of the second hypothesis, namely, that the differences are due to the different relationships of the two stresses to the experimental task. Briefly, the process postulated is somewhat like this. The failure stress served to thwart a motive which was directly involved in the experimental task. Presumably, this produced interference with performance. In the case of the fear stress the thwarted motive bore no direct relation to the one under which *S*s were working and therefore was less likely to interfere with performance.

Another important question that presents itself at this point concerns the adequacy of the experimental task for the purpose of observing intellectual performance under two stressful situations. The problems in long division fulfill several criteria considered essential for this type of task, namely, high test-retest reliability, the need for concentration and the use of judgment, and also the relatively long span of attention for completion of a problem. A possible weakness of this type of task is that it represents performance that is so stable for a given individual that it is not readily affected by experimental conditions. Also, arithmetic problems differ in some respects from other types of intellectual tasks. For example, the judgment required is either right or wrong—there is no middle ground which would permit the observation of degree of error or qualitative changes in performance. There is also the possibility of immediate verification of the judgment made and subsequent self-correction. For example, if *S* decides that the divisor goes into the dividend four times, whereas it actually goes in five times, he discovers this as soon as he performs the necessary multiplication and subtraction. This does, of course, slow down his work. Whether the decrement is produced by an over-all slowing down of performance or by an increase of poor judgments which need correction, or by still other factors cannot be determined. The data in this experiment fail to show an increase in errors, but do show a decrease in performance. Because of the self-correction possible with this task, we cannot be sure that this gives us a complete picture of what actually happened. Thus differences in the effects of the two stress situations could have been obscured. Lindsley (8) and Lazarus and Eriksen (7) reported an increase in the number of errors under stress on tasks where no self-correction was possible.

In these respects the problems in long division are more limited than other types of tasks which can reveal the number and types of errors made. Before engaging in any generalizations regarding the effects of various types of stress upon intellectual performance many more aspects of the latter will need to be investigated.

The question of the relationship between *S*s' level of anxiety and performance under stress deserves some comment. Despite the negative findings obtained in this study and despite the weaknesses of an anxiety score derived from a paper-and-pencil test, the author is inclined to believe that positive results would have been obtained if the experiment had been differently designed. This belief is based on the following facts.

First, the correlation coefficients between the level of manifest anxiety and difference score, though not significantly different from zero, are negative for the three stressed groups and positive for the two control groups. The probability of this being a chance occurrence is less than .05.

Second, when *Ss* in each experimental group are divided into two subgroups, according to whether they are above or below the median anxiety score, the differences in performance under stress between these subgroups, within each stressed group, are of the same order of magnitude as the differences between the stressed and control groups in the original data. Furthermore, these differences do not exist within the control groups. It would have been desirable to match *Ss* originally on the level of manifest anxiety as well as on initial performance. If this had been done, a more sensitive test of significance could have been performed than is afforded by a correlation coefficient between anxiety level and difference score. Such a test might have revealed that anxiety plays a significant role in performance under stress.

Summary

The effect of two types of stress, failure and fear, upon performance, and the role played by the level of manifest anxiety of *Ss* were observed. The task was an arithmetic test which consisted of problems in long division. The *Ss* were 305 high school students, divided into five groups which were subjected to the following experimental treatments: I. failure; II. failure and fear; III. success and fear; IV. success; V. control.

An analysis of covariance yielded the following results:

1. The effect of failure was to depress performance.

2. Fear produced no significant change in performance.

3. Individual group comparisons with the control group demonstrated that experimental results vary depending on whether comparisons are made with a success group or with a group under neutral motivation.

4. Anxiety produced no significant change in performance. However, the correlation coefficients between level of anxiety and difference score were negative for the three stressed groups and positive for the two nonstressed groups. An improved experimental design is suggested for a more effective study of this variable.

The hypothesis is proposed that the difference in effects produced by failure and fear may be due to the relation between the stress and the experimental task.

References

1. Alper, T. G. Memory for completed and incompleted tasks as a function of personality: an analysis of group data. *J. abnorm. soc. Psychol.*, 1946, **41**, 403–421.
2. Deese, J., Lazarus, R. S., & Keenan, J. Anxiety, anxiety reduction, and stress in learning. *J. exp. Psychol.*, 1953, **46**, 55–60.
3. Glixman, A. F. Recall of completed and incompleted activities under varying degrees of stress. *J. exp. Psychol.*, 1949, **39**, 281–295.
4. Goulden, C. H. *Methods of statistical analysis.* New York: Wiley, 1952.
5. Lazarus, R. S., Deese, J., & Hamilton, R. Anxiety and stress in learning: the role of intraserial duplication. *J. exp. Psychol.*, 1953, **47**, 111–114.
6. Lazarus, R. S., Deese, J., & Osler, S. F. The effects of psychological stress upon performance. *Psychol. Bull.*, 1952, **49**, 293–317.
7. Lazarus, R. S., & Eriksen, C. W. Effects of failure stress upon skilled performance. *J. exp. Psychol.*, 1952, **43**, 100–105.
8. Lindsley, D. B. *A study of performance under speed stress.* OSRD, 1945, Publ. Bd. No. 1838. Washington, D. C.: U. S. Dept. of Commerce, 1946.
9. Lucas, J. D. The interactive effects of anxiety, failure and intraserial duplication. *Amer. J. Psychol.*, 1952, **55**, 59–66.
10. McKinney, F. Certain emotional factors in learning and efficiency. *J. gen. Psychol.*, 1933, **9**, 101–116.
11. Montague, E. K. The role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, **45**, 91–96.
12. Postman, L., & Bruner, J. S. Perception under stress. *Psychol. Rev.*, 1948, **55**, 314–323.
13. Stopol, M. S. Rorschach performance in relation to two types of stress. *J. consult. Psychol.*, 1954, **18**, 11–15.
14. Sullivan, E. B. Attitude in relation to learning. *Psychol. Monogr.*, 1927, **36**, No. 3 (Whole No. 169).
15. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
16. Taylor, J. A., & Spence, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, **44**, 61–64.
17. Williams, M. An experimental study of intellectual control under stress and associated Rorschach factors. *J. consult. Psychol.*, 1947, **11**, 21–29.

(Received May 25, 1953)

Journal of Experimental Psychology
Vol. 47, No. 2, 1954

FUNCTIONAL FIXEDNESS AS RELATED TO ELAPSED TIME AND TO SET[1]

ROBERT E. ADAMSON AND DONALD W. TAYLOR
Carleton College *Stanford University*

Duncker (3) originally proposed the hypothesis that problem solving may be inhibited by the functional fixedness of solution objects. That is, owing to previous use of the object with a function dissimilar to that demanded by the present problem, *S* is inhibited in discovering the appropriate new use. In a series of six experiments he obtained results supporting the hypothesis. An earlier paper (1) reported a repetition of three of Duncker's experiments, using a larger *N* and with more carefully specified experimental conditions. The results essentially confirmed those of Duncker, and, together with evidence from a study by Birch and Rabinowitz (2), offer convincing proof of the existence of functional fixedness. The present study follows two lines of investigation designed to clarify the nature of the phenomenon: (*a*) determination of conditions influencing its occurrence, and (*b*) exploration of its relation to other kinds of set in problem solving.

It would appear that the fixedness of an object results from its association with a particular function. As a result of this association it becomes more difficult to utilize the object with a new function. One inference from this is that anything which weakens the association between an object and a specific function will lessen the amount of functional fixedness for that object. Duncker has stated, along these lines, that "separation of the two functions F_1 and F_2 in two independent problem situations about eliminates the disturbing functional fixedness of the crucial object" (3, p. 96). But the Birch and Rabinowitz study (2) incorporated such a change of context and still demonstrated fixedness for the solution objects. One factor which might clearly be expected to weaken the association between an object and a specific function would be lapse of time following the use of that object with that function. This expectation suggests the hypothesis that functional fixedness will decrease with increasing time following initial usage of the object. Experiment I was designed to test this hypothesis.

Studies by Luchins (5, 6) and Guetzkow (4) have been concerned with the operation of set in problem solving, where set is defined as the continued attempt to use a previously successful method in problems where the method is no longer adequate. Both functional fixedness and set thus defined involve the limiting effect of past experience; in both the variability of approach to a problem is reduced. This similarity suggests that individuals susceptible to functional fixedness would also tend to be susceptible to such set. Guetzkow (4) has distinguished susceptibility to set from the inability to overcome set. Experiment II was designed to test the hypothesis that susceptibility to func-

[1] These experiments were part of a series being carried out under Project NR 150-104 and supported by Contract Nonr 225 (02) between Stanford University and the Office of Naval Research. The first author designed and conducted the experiments reported here. The second author carried out the statistical analysis of the data and prepared the present report. Work on the contract is now under the general direction of the second author. Permission is granted for reproduction, translation, publication, use, and disposal of this article in whole or in part by or for the United States Government.

tional fixedness is positively related to susceptibility to set and to inability to overcome set.

Experiment I: The Influence of Elapsed Time on Functional Fixedness

Procedure.—Five experimental groups, including 16, 16, 14, 11, and 10 *Ss*, respectively, participated in the first experiment. The *Ss* were randomly assigned to the various groups and all were undergraduate students in general psychology at Stanford University.

The procedure employed was a modification of that introduced by Birch and Rabinowitz (2), a procedure which allows the introduction of a varying time interval between the initial usage of an object and its use in the test problem. The members of the experimental group were first "given practice" in completing electric circuits with either a switch or a relay. In the present study, each *S* in the experimental groups was initially given a "speed test" requiring the construction of an electric circuit following a drawing. For half of the *Ss* in each group the circuit incorporated a microswitch; the other half completed a circuit using a small relay.

Between this initial task and the test problem, a period of uncontrolled activity occurred, the length of time being different for each of the five groups (see Fig. 1). Following this interval all *Ss* were given Maier's "two-string" problem (7). This problem consists of tying together the ends of two strings which are too far apart to be reached simultaneously unless one of the strings is moored or swinging. Since the method of presenting the problem prevented the mooring of either string, the only solution was to swing one of them while procuring the other. The oscillations of either string dampened too rapidly for it to swing by itself, and a weight attached to the end of the string was necessary. The *S* was presented with a number of possible solution objects, among them being the switch and the relay, the only two heavy enough to serve as a pendulum weight for the string. The purpose was not to determine whether *S* could solve the problem; rather it was to determine which of the two objects would be used to solve the problem, the one used in the initial task or the one not used in the initial task. If *S* had not solved the problem at the end of 3 min., *E* brushed against one of the strings so as to start it swinging. No *S* failed to solve the problem within the time allotted.

Results.—In the initial task, half of the *Ss* in each experimental group had been given the switch and half the relay to use in completing the circuit. Therefore, in the two-string problem the null hypothesis would be that by chance 50% of each group would use the object previously used and 50% the one not previously used.[2] Occurrence of functional fixedness would result in more than 50% of a group choosing the object not used in the initial task. The results obtained are presented in Fig. 1.

Of the 16 *Ss* in each of the first two experimental groups, 11 (69%) used in solving the two-string problem the object not used in the initial task. If these two groups are combined and the significance of the difference between the obtained 69% and the

[2] In choosing the microswitch and relay to be used in the experiment, the attempt was made to select two objects whose physical characteristics would make them equally appropriate for use as pendulum weights. A group of 16 *Ss* were given the two-string problem, but not the initial task involving completion of the circuit using the switch or relay. In solving the two-string problem, eight *Ss* chose the switch and eight the relay.

It should be noted, however, that the null hypothesis does not depend on this finding that no preference was exhibited for one of the objects. Since in the initial task half of each experimental group received the microswitch and half the relay, the null hypothesis in the two-string problem would be that half would choose and half avoid the object previously used even though the group as a whole exhibited a preference for the microswitch (or relay) as a pendulum weight. Suppose, for example, that of 100 *Ss* in an experimental group, 70 chose the microswitch and 30 the relay. The null hypothesis would be that of the 70, 35 would be from the group of 50 given the switch in the initial task and 35 from the group of 50 given the relay. Similarly, of the 30, 15 would be from the group given the switch and 15 from the group given the relay. Thus, even though a definite preference is exhibited for the microswitch, the null hypothesis is that 35 + 15 or 50 will choose and 35 + 15 will avoid the object previously used. It should be noted further, however, that in the case of such a preference, a maximum of 80 of the 100 *Ss* could choose the object not previously used. The two-string problem will be most sensitive as a measure of functional fixedness if the two objects are equally appropriate for use as pendulum weights.

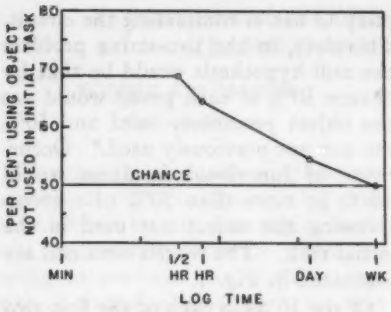


FIG. 1. Functional fixedness as a function of log time

expected 50% computed, a critical ratio of 2.13 is obtained which with a one-tail test is significant at the .02 level. Experience in the initial task did produce functional fixedness.

The experiment was designed, however, to test the more specific hypothesis that functional fixedness decreases with increasing time. Figure 1 shows that the five groups rank in the order predicted, with the exception that the first two groups are tied. Computation of tau for the correlation between the predicted and the obtained rank orders yields a value significant at the .02 level (8). Thus, the hypothesis is confirmed.

Discussion.—The conclusion that functional fixedness decreases as a function of time is not to be interpreted as assigning causal significance to time as such. This decrease is probably best interpreted in terms of the concept of retroactive inhibition. It appears likely that the activity during the period between the initial task and the two-string problem tends to interfere with the retention of the association formed during the initial task. The longer the interval, the larger the amount of intervening activity, the greater the degree of retroactive inhibition, and hence the smaller the amount of functional fixedness obtained. The nature of the activity during the intervening interval was uncontrolled in the present experiment. It would be expected that the decrease in functional fixedness with time would be a function of the kind of intervening activity. Additional experiments will be necessary to test this hypothesis.

EXPERIMENT II: THE RELATION BETWEEN FUNCTIONAL FIXEDNESS AND SET

Procedure.—All except one of the 46 *S*s from the first three experimental groups in Exp. I were also given a series of 11 problems devised by Luchins (5). In each problem, the task is to explain how a stipulated quantity of water could be obtained by using jars of given sizes; e.g., given an empty 14-qt. jar, an empty 36-qt. jar, and an empty 8-qt. jar; measure 6 qt. water. Problem 1 involves only two jars and is given to illustrate the nature of the task. Problems 2 to 6 are all solved in the same way: first fill the second jar, then subtract from it the amount necessary to fill the first jar once and the third jar twice. Problems 7, 8, 10, and 11 can also be solved in this way or, more directly, in one step which does not involve the second jar. Problem 9 can be solved only by the more direct method. The time allowed for each problem was 1.5 min.

Luchins (6) has pointed out that two types of behavioral rigidity should be distinguished in performance on the water-jar problems. Use of the indirect instead of the simpler direct method in solving Problems 7 and 8 may be taken as evidence of one kind of behavioral rigidity. Since use of the indirect method, even though less efficient, does result in solution, designating such behavior as rigid may be questioned. In any case, this kind of rigidity should be distinguished from that involved in Problem 9 where failure to use the direct method results in failure to solve the problem.

Guetzkow (4) has used somewhat different terminology in making the same distinction. Use of the indirect method in solving Problems 7 and 8 is taken as evidence of susceptibility to set. Failure to solve Problem 9 is evidence of inability to overcome set.

Results.—The first hypothesis to be tested was that susceptibility to functional fixedness is related to susceptibility to set. Data presented in Table 1, however, show that no significant relation was obtained between the choice of solution objects in the two-string problem and the method used in solving Problems 7 and 8.

In contrast, the results clearly support the second hypothesis that sus-

TABLE 1

χ^2 CONTINGENCY TABLES: FUNCTIONAL FIXEDNESS, SUSCEPTIBILITY TO SET, AND INABILITY TO OVERCOME SET

Luchins Problems	Nature of Solution	Two-String Problem		χ^2	p*
		Object Previously Used	Object Not Previously Used		
7 and 8	Both indirectly	10	23	.13	—
	One or both directly	5	7		
9	Failed	3	20	6.95	<.005
	Solved	12	10		
10 and 11	Both indirectly	4	17	2.51	<.06
	One or both directly	11	13		

* Since the direction of the relation was predicted, a one-tail test was employed, i.e., the square root of χ^2 was interpreted by means of the normal curve table. Yates's correction was employed in all computations of χ^2 in this article.

ceptibility to functional fixedness is related to inability to overcome set. As shown in Table 1, those who solve the two-string problem by using the object not used in the initial task are more likely to fail Problem 9. The argument may be plausibly made that this analysis should include only those *Ss* who solved Problems 7 and 8 indirectly: if performance on Problem 9 is to be a criterion of ability to overcome set, its use must be limited to those *Ss* for whom the existence of set has been demonstrated on Problems 7 and 8. A plausible counterargument would be that those *Ss* who solve Problems 7 and 8 directly have done so in the face of the same set-inducing experience on Problems 2 to 6 and have thereby demonstrated their ability to overcome set. In any case, the analogous fourfold table including only the 33 *Ss* who solved Problems 7 and 8 indirectly yielded a χ^2 of 11.61, significant at the .0004 level.

The relation between choice of solution objects in the two-string problem and the method used in solving Problems 10 and 11 approaches significance, p being .06. Actually, if the analysis is limited to the 41 *Ss* who solved Problems 10 and 11 both directly or both indirectly, a χ^2 of 3.10 and p of .04 are obtained.

Discussion.—The distinction between two types of behavioral rigidity, or in other terms between susceptibility to set and inability to overcome set, is apparently an important one. Luchins (6) found that behavioral rigidity measured by performance on problems like Problem 9 is positively related to concreteness of thinking as measured by the Similarities subtest of the Wechsler-Bellevue Intelligence Scale. No significant relation was found between behavioral rigidity of the kind measured by performance on problems like Problems 7 and 8 and concreteness of thinking. Guetzkow (4) obtained a significant difference between men and women in inability to overcome set, but found no such difference in susceptibility to set.

The present finding that susceptibility to functional fixedness is related to inability to overcome set, but not to susceptibility to set confirms the importance of the distinction. Continuing to use an effective indirect method, even though a more direct method is available, does not have the same significance psychologically as does continued use when the method is inappropriate.

The fact that the relation between susceptibility to fixedness and performance on Problems 10 and 11 approaches

significance is not in conflict with the statement just made. It is highly probable that performance on Problems 10 and 11 was affected by performance on Problem 9. Those who solve Problem 9 become aware of the direct method and may be expected to use it in solving Problems 10 and 11; in fact, of the 22 who solved Problem 9, only 1 person solved both 10 and 11 indirectly. The obtained relation between fixedness and performance on Problems 10 and 11 may be best interpreted as owing to the effects of experience with Problem 9.

Attention should be called to an important limitation of the measure of functional fixedness employed here. The 30 of the 45 *S*s who in the two-string problem employed the object not previously used were in the analyses of the data of Exp. II all classified as susceptible to functional fixedness. This classification is open to question. As pointed out in the analysis of Exp. I, the null hypothesis would be that by chance 50% use the object not previously used. Hence the 30 *S*s include those who by chance chose the object not previously used as well as those susceptible to functional fixedness. Unfortunately, there is no way to discriminate between them. For the purposes of Exp. II, therefore, the measure of functional fixedness employed must be regarded as a relatively insensitive one. The fact that in spite of this insensitivity a highly significant relation between functional fixedness thus measured and inability to overcome set was obtained deserves emphasis. This insensitivity must also be kept in mind in interpreting the fact that no significant relation was obtained between functional fixedness and susceptibility to set.

Summary

Five groups including a total of 67 *S*s participated in Exp. I designed to test the hypothesis that functional fixedness decreases as a function of time. All *S*s were first given the task of constructing an electric circuit. For half of the *S*s in each group, the circuit employed a microswitch and for the other half a small relay. Following an interval of uncontrolled activity, all *S*s were given Maier's two-string problem. Among the objects available for possible use as a pendulum weight, the only two which would work were the microswitch and the relay. The purpose was to determine which of the two objects would be used in solving the problem. The length of the interval between the initial task and the two-string problems was varied, being 1 min., ½ hr., 1 hr., 1 day, and 1 week for the five groups, respectively. Functional fixedness, as measured by the percentage of the *S*s using as a pendulum weight the object not used in the initial task, did decrease as a function of time. The relation was significant at the .02 level.

In Exp. II, concerned with the relation of functional fixedness to set, 45 *S*s from the first three groups in Exp. I were given a Luchins series of water-jar problems. Problems 2 to 6 are all solved by the same indirect method. Performance on Problems 7 and 8, which can be solved either by this indirect method or by a more direct method, was used as the criterion of susceptibility to set. Performance on Problem 9, which can be solved only by a direct method, served as the criterion of inability to overcome set. Susceptibility to functional fixedness, as defined by the Maier two-string problem, was found to have a significant relation to inability to overcome set, but no relation to susceptibility to set, as defined by the Luchins problem series.

References

1. Adamson, R. E. Functional fixedness as related to problem solving: a repetition of three experiments. *J. exp. Psychol.*, 1952, **44**, 288–291.
2. Birch, H. G., & Rabinowitz, H. S. The negative effect of previous experience on productive thinking. *J. exp. Psychol.*, 1951, **41**, 121–125.
3. Duncker, K. On problem-solving. *Psychol. Monogr.*, 1945, **58**, No. 5 (Whole No. 270).
4. Guetzkow, H. An analysis of the operation of set in problem-solving behavior. *J. gen. Psychol.*, 1951, **45**, 219–244.
5. Luchins, A. S. Mechanization in problem solving. *Psychol. Monogr.*, 1942, **54**, No. 6 (Whole No. 248).
6. Luchins, A. S. The Einstellung test of rigidity: its relation to concreteness of thinking. *J. consult. Psychol.*, 1951, **15**, 303–310.
7. Maier, N. R. F. Reasoning in humans. I. On direction. *J. comp. Psychol.*, 1930, **10**, 115–143.
8. Stilliro, G. P. The distribution of Kendall's coefficient of rank correlation in rankings containing ties. *Biometrika*, 1947, **34**, 36–40.

(Received for early publication October 22, 1953)

Journal of Experimental Psychology
Vol. 47, No. 2, 1954

THE RELATION OF ANXIETY TO DIFFERENTIAL EYELID CONDITIONING[1]

KENNETH W. SPENCE AND I. E. FARBER

State University of Iowa

A number of recent studies from the Iowa laboratory have attempted to explore and extend the behavioral theory of motivation developed originally by Hull on the basis of simple conditioning studies (12). According to this theory the various drive-producing conditions in the environment, e.g., periodic food and water deficiencies, noxious stimulation, conditioned fear cues, etc., are assumed to combine in their action to determine the total drive level of the organism at any moment. This total effective drive strength, represented by *D*, is one of the important intervening variables determining response strength in Hull's formulation.

That response strength, at least in very simple learning situations, is a function of the level of drive as determined by one or more of the drive-producing conditions had been shown in a number of earlier studies (22, 23, 28). Indeed, Hull's initial formulation of his drive theory was fashioned on the basis of these data. Subsequent investigations of the need or drive combination aspects of the theory have provided encouraging evidence in its support. Thus the studies of Kendler (14), Webb (29), and Amsel (1) all showed that under certain conditions different needs combine with each other to produce greater response strength. More recently a series of studies involving classical eyelid conditioning of so-called anxious and nonanxious human *Ss* has been offered in further support of the theory (25, 26, 27).

These latter studies took their start from the theoretical paper by Mowrer (20) and the findings of such investigators as Miller (19) and Brown and Jacobs (3) that conditioned pain or fear provided a state or condition (level of *D*?) that led to striving behavior and the occurrence of learning. Also, such experiments as those of Amsel (1) and Brown, Kalish, and Farber (4) had indicated that the presence in their test situation of cues that had been conditioned to arouse emotional (fear) activity increased the response strength. An obvious implication of these findings for Hull's theory is that level of *D* is a function of the level of emotionality present in the organism.

Accepting this relation as a working hypothesis, Taylor (27) attempted to vary the level of emotionality in human *Ss* by selecting two groups differing in emotionality as determined by a personality inventory. The latter consisted of items which clinical psychologists had judged to be symptomatic of manifest anxiety (i.e., overt signs of fear, worry, emotionality, etc.). Selecting *Ss* from the upper and lower extremes of the distribution of scores on this scale, Taylor predicted that those from the high emotionality end (high *D*) would exhibit a higher level of performance in the conditioning situation than would *Ss* from the low end of the scale. The

[1] This study was carried out as part of a project concerned with the influence of motivation on performance in learning under Contract N9onr-93802, Project NR 154-107 between the State University of Iowa and the Office of Naval Research. The conditioning data were collected by Elaine Taylor.

results of this first conditioning study confirmed Taylor's hypothesis, as have two subsequent conditioning experiments from our laboratory (**25, 26**). These experiments suggest, then, that *Ss* who admit to having a lot of emotional symptoms have a relatively high D level and hence respond at a relatively high level in a simple learning situation such as classical conditioning.

The prediction that a high D level leads to a high level of response is based, of course, on Hull's theory for simple, noncompetitional response situations that

$$R = f(E) = f(H, D)$$

In this type of situation, it will be noted, there are no competing responses but only a single S–R tendency. Thus, in classical conditioning there is, by virtue of the control of the stimulus conditions, a single highly dominant response. As training proceeds, the CS acquires greater and greater habit strength (H) for this response. The excitatory strength (E) of the CS to evoke the response depends, it is assumed, on the product of this H and the level of D, i.e., $E = H \times D$. The higher the value of D, the greater E will be and hence the greater will be the response strength.

The present study extends the testing of our hypothesis concerning the relation of Taylor Anxiety Scale scores to level of D, and the role of the latter in classical conditioning to the differential conditioning situation. A number of investigators (**8, 9**) have shown that in differential conditioning, in which one stimulus, designated S_+, is always reinforced by the UCS and a second similar stimulus, designated S_-, is never reinforced, an initial period of generalization occurs in which the CR develops to both stimuli. With continued differential training the CR to S_- eventually disappears.

In terms of behavior theory the occurrence of the response to S_- occurs by virtue of generalization to S_- of the habit strength (H) developed to S_+. The magnitude of this generalized habit strength ($\overline{H}$) is assumed to be a function of (*a*) the strength of H, i.e., of the number of reinforced trials to S_+, and (*b*) the magnitude of the difference between the two stimuli, S_+ and S_-. Response strength gradients confirmatory of such an assumed gradient of generalization of H with stimulus differences have been obtained in a number of studies (**2, 6, 11**). While less evidence is available as to how generalized habit strength ($\overline{H}$) develops with conditioning trials to the original CS, a recent study reported by Margolius (**16**) suggests that the relation is a positive monotonic one. Fortunately, for our present purposes knowledge of the precise nature of these functions is not necessary. Rather our interest is in the effects of different levels of D on (*a*) the strength of the responses to the two stimuli, positive and negative, and (*b*) on the magnitude of the difference between these strengths, i.e., on discrimination of the two stimuli.

According to our theory, the excitatory strengths (E_+ and E_-) of the positive and negative stimuli should be higher in the case of the anxious (high D) group than for the nonanxious (low D) group. That is

$$\text{Anxious } E_+ = H \times D_{\text{anx.}}$$
$$> \text{Nonanxious } E_+ = H \times D_{\text{nonanx.}}$$
$$\text{Anxious } E_- = \overline{H} \times D_{\text{anx.}}$$
$$> \text{Nonanxious } E_- = \overline{H} \times D_{\text{nonanx.}}$$

Furthermore, the difference between E_+ and E_- should also be a function of D. Thus, if

$$E_+ = H \times D$$
$$\text{and } E_- = \overline{H} \times D,$$
$$\text{then } E_+ - E_- = D(H - \overline{H}).$$

Thus, in terms of excitatory strength, our theory would lead us to expect nonanxious *Ss* to exhibit less discrimination of the positive and negative stimuli than anxious *Ss*.[2]

It is important to note that the above derivation refers to excitatory potential (E) and not to the response measure, percentage or frequency of CR's (R_p). As one of the writers (24) has shown, Hull's theory of oscillation ($_sO_R$) leads to the derivation of a necessary relation between R_p and E, namely, that of the normal probability integral. By means of this relation it is possible, as Hull has done in his recent book (13), to transform measures of R_p into measures of E. In the experiments to be reported all measures are in terms of E.[3] Two separate experiments differing slightly in procedure were conducted to test these theoretical deductions.

Method

Experiment I

Subjects.—The *Ss* were selected from a course in introductory psychology on the basis of scores made on the Taylor Anxiety Scale (27). The anxious group consisted of 18 *Ss*, 10 men and 8 women, whose scores on the scale (21–32) were in the upper 20% of scores made by Iowa students, while the nonanxious group was composed of 18 *Ss*, 10 men and 8 women, whose scores (1–6) fell in the lower 20%. Sixteen other *Ss* were run in the experiment; 14 of these, 6 from the anxious group and 8 from the nonanxious group, were eliminated because of failure to meet a criterion that had been used in previous studies (25, 26) as a means of excluding *Ss* who gave a high incidence of voluntary responses. As in the earlier experiments, the occurrence of more than 50% CR's with latencies less than 300 msec. was accepted as an indication that *S* was a voluntary responder.[4] Two *Ss*, one from each anxiety category, were eliminated because they gave no CR's to the positive stimulus.

Apparatus and method of recording.—The equipment for recording eyelid closure and presenting the UCS was identical with that used in a previous study (26). The positive CS (S_+) was a 500-cycle tone produced by a loud-speaker, driven by a Hewlett-Packard oscillator. The negative CS (S_-) produced by the same means was a 5,000-cycle tone. The sound pressure levels of the tones at the position of the *S*'s head, as measured by means of a General Radio sound level meter, Type 759-B, were 66 db for S_+ and 49 db for S_-. The duration of both tones on each trial was 550 msec., with the UCS occurring 500 msec. after the onset in the case of the positive CS. The UCS was omitted on trials on which the negative CS was presented. The duration of the UCS, an air puff of 1.0 lb./sq. in. applied to the right eye, was limited to 50 msec. by means of a 110-v., 60-cycle, AC-operated solenoid valve controlled by an electronic timer.

A CR was recorded whenever the record showed a deflection of 1 mm. or more in the interval 150 to 500 msec. following the onset of the CS. Responses with a latency of less than 150 msec. were classified as original responses to the CS and were not included in the data.

Differential conditioning procedure.—Following the reading of instructions, each *S* received three presentations of the positive CS alone. A single presentation of the UCS was then given and the response of the eyelid recorded for the ensuing 40 sec. Immediately following these preliminary trials each *S* received 100 trials, 50 of which involved the presentation of the positive cue (S_+) with the UCS and 50 of which involved

[2] There are, of course, a number of other postulates in our S–R theory that we have assumed the reader will be aware of on the basis of his knowledge of Hullian theory. One to which attention should perhaps be drawn is that the internal drive stimuli (S_D) of the two groups of *Ss* are assumed to have relatively negligible and equal habit loadings for the CR. Indeed, in view of the fact that the onset of these stimuli is quite uncorrelated with the occurrence of the eyelid response, there is little reason to suspect that any habit strength whatever is acquired by these stimulus cues.

[3] Such numerical transformation of measures for purposes of testing a theory is common practice in the more advanced fields of science. Mueller (21) has given an excellent account of such transformations and their various uses in the field of psychology. Consideration of the normal integral relation between R_p and E will show that the same prediction of the effects of different drive levels on differentiation will hold for low and medium levels of the measure R_p as holds for E, but that it will not necessarily hold once the frequency of response to S_+ becomes relatively high.

[4] The percentage of such "voluntary responders" (27%) in this differential conditioning situation was considerably higher than in our previous conditioning studies (25, 26), in which approximately only 10% of the *Ss* fell into this category.

TABLE 1

EXPERIMENT I: EXCITATORY STRENGTHS OF POSITIVE AND NEGATIVE STIMULI ON TRIALS 61–100 FOR ANXIOUS AND NONANXIOUS *Ss*

Level of Drive	Men (N = 20)						Women (N = 16)					
	S_+		S_-		Diff.		S_+		S_-		Diff.	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Anxious (N = 18)	2.88	.49	1.58	.73	1.29	.57	3.51	.76	2.00	1.02	1.51	1.38
Nonanxious (N = 18)	2.34	.69	1.22	.69	1.12	1.02	2.72	1.00	2.12	.60	.61	.50

the negative cue (S_-) without the UCS. The order of presentation of S_+ and S_- was prearranged so that no more than three reinforced nor more than two nonreinforced trials occurred in succession. Intertrial intervals of 15, 20, and 25 sec., averaging 20 sec., and arranged according to a fixed schedule, were used. At the end of the session all *Ss* were questioned as to their understanding of the purpose of the experiment and warned not to discuss the experiment with other members of the class. Approximately one-third of the men and an equal proportion of the women were able to state that the experiment was concerned with conditioning. However, this knowledge did not affect the conditioning scores in any significant manner.

Experiment II

Subjects.—One hundred and two *Ss*, 37 men and 65 women, from a course in introductory psychology served in the experiment. The data of nine additional *Ss*, four anxious and five nonanxious, were excluded because of a high incidence of voluntary responses. Five other *Ss*, two from the anxious category and three from the nonanxious, were eliminated because of failure to give any CR's to the positive CS.

A somewhat different criterion of anxiety from that employed in previous studies was used in this experiment. In addition to the Taylor Anxiety Scale, scores were also available on a forced-choice form of the Scale developed by Heineman (7). Scores on this latter scale have been shown to be relatively unsusceptible to distortion resulting from attempts to minimize anxiety. The *Ss* for the present experiment were selected on the basis of scores falling in the upper and lower 20 percentiles of either the Taylor Scale or the Heineman forced-choice form. This dual criterion allowed for the inclusion of a certain number of *Ss* who fell below the 80th percentile score and above the 20th percentile score of the Taylor Scale. The chief reason for using this dual criterion was to provide a larger number of *Ss*.

Apparatus and procedure.—The apparatus and method of recording data were exactly the same as in Exp. I. The only difference in procedure was that instead of omitting the UCS following the negative CS, a puff of air was given 2,500 msec. after the onset of S_-. The duration of the latter was extended to 2,550 msec. A CR to S_- was defined, as before, as the occurrence of an eyelid closure of 1 mm. or more in the period 150 to 500 msec. following the onset of the stimulus. Previous studies from this laboratory (17, 18, 25) have shown that a CR established at an interval of 500 msec. extinguishes when the interval between the CS and UCS is extended to 2,500 msec.

RESULTS

Experiment I

According to the present theory the excitatory strengths of the corresponding cue stimuli should be higher for the anxious *Ss* than in the case of the nonanxious *Ss*. Furthermore, there should be a larger difference between the excitatory strengths of the positive and negative stimuli in the case of the anxious *Ss* than for the nonanxious *Ss*. Table 1 presents the data relevant to these implications, based on the last 40 trials (i.e., Trials 61–100).[5] The values in the table are in terms of excitatory potential (E), as determined from the percentage of responses made to the stimuli by means of Table 34 in Hull's *Principles of Behavior* (12).

It will be observed that five of the six differences were in the predicted direction. Thus in the case of the positive stimulus both the anxious

[5] At this stage of training the values of H and $\bar{H}$ should be maximal, thus providing for maximal differences in E for different D levels.

TABLE 2

EXPERIMENT I: SUMMARY OF ANALYSIS OF VARIANCE OF *E* VALUES BASED ON TRIALS 61–100 OF ANXIOUS AND NONANXIOUS *Ss*

Source	*df*	*MS*	*F*	*p*
Anxiety	1	2.936	2.50	.12
Sex	1	5.896	5.02	<.05
Anxiety × Sex	1	0.708	—	
Error (*b*)	32	1.174		
Stimuli	1	23.461	48.46	<.001
Stimuli × Anxiety	1	1.090	2.25	.14
Stimuli × Sex	1	.096	—	
Stimuli × Anxiety × Sex	1	.613	1.26	>.20
Error (*w*)	32	.484		

men and anxious women exhibited a stronger excitatory strength than the nonanxious groups. Likewise, in the case of the differences between the excitatory strengths of the cue stimuli, the values for the anxious groups are larger than those for the nonanxious groups in both sexes. Only in the case of the negative stimulus was a difference in the direction opposite to that predicted obtained. Here it will be seen that the nonanxious women had a higher *E* value than the anxious women. The men showed the predicted result.

A summary of an analysis of variance (Lindquist Type III design [15]) based upon these data is presented in Table 2. It may be seen that the over-all difference between the excitatory tendencies to S_+ and to S_- was highly significant and that the over-all difference between the two sexes was also reliable. The main effect of anxiety, however, was significant at only the .12 level of confidence. Also the interaction between stimulus cue and anxiety was not reliable ($p = .14$).

As in previous studies the anxious *Ss* showed a higher level of response strength to the positive CS than did the nonanxious *Ss*. The obtained *t*'s were 1.91 and 1.88, for men and women, respectively, with $p < .05$ for a single-tailed hypothesis in both instances. The differences between the response levels of the anxious and nonanxious *Ss* on the negative CS were not significant. The predicted differences between S_+ and S_- approached, but did not attain, significance. The *t* for men was 1.30 ($p < .11$) and for women 1.62 ($p < .07$).[6]

TABLE 4

EXPERIMENT II: SUMMARY OF ANALYSIS OF VARIANCE OF *E* VALUES BASED ON TRIALS 61–100 FOR ANXIOUS AND NONANXIOUS WOMEN

Source	*df*	*MS*	*F*	*p*
Anxiecy	1	6.9745	5.12	<.05
Error (*b*)	63	1.3604		
Stimuli	1	10.2481	36.45	<.001
Anxiety × Stimuli	1	.2961	1.05	>.20
Error (*w*)	63	.2811		

[6] In all instances in which the effects of anxiety are in the predicted direction, the relevant *t* tests of significance are based upon a single-tailed hypothesis. All levels of significance greater than .05 were computed in accordance with a procedure described by Burke (5). Analyses of the results in terms of percentages of CR's presented essentially the same picture as the *E* values. There was a general tendency for the differences involving percentages to be slightly more significant than the corresponding *E* differences.

TABLE 3

EXPERIMENT II: EXCITATORY STRENGTHS OF POSITIVE AND NEGATIVE STIMULI ON TRIALS 61–100 FOR ANXIOUS AND NONANXIOUS WOMEN

Level of Drive	S_+		S_-		Diff.	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Anxious ($N = 40$)	3.00	.97	2.36	.84	.64	.75
Nonanxious ($N = 25$)	2.42	.71	1.98	.97	.44	.72

Experiment II

The results for men and women in this experiment were analyzed separately because of lack of proportionality of the sexes in the anxious and nonanxious groups. Table 3 indicates that the anxious women showed a higher response level to both the positive and negative stimuli than the nonanxious women, and that the difference between the excitatory strengths of the two stimuli was also greater for the anxious *Ss* than for the nonanxious *Ss*.

Table 4 summarizes the results of an analysis of variance of these data. It will be noted that the *F* for anxiety was 5.12, which is significant at better than the .05 level of confidence. The interaction between anxiety and stimuli was not significant, however. As in Exp. I the response strength to the positive stimulus was significantly greater in the case of the anxious women than the nonanxious women. The *t* was 2.52, which is significant at better than the .01 level of confidence. The *t* obtained for the negative stimulus was 1.65, which is significant at approximately the .05 level of confidence. The difference between the *E*'s of the cue stimuli was, as predicted, in favor of the anxious group, but the obtained *t* of 1.05 was not significant ($p = .15$).

Table 5 presents the data for the men. While the differences between the anxious and nonanxious *Ss* again were all in the predicted direction, none of them was significant.

Discussion

While not conclusive, the results of these two experiments tend to be in agreement with the implications of our theory concerning the relation of anxiety to level of *D* and the relation, in turn, of *D* to response strength in classical conditioning. Thus, of the 12 predicted differences involving anxiety in the two experiments, 11 were in the expected direction. Of these, however, only four were significant at the .05 level or better; three others approached significance, while four were quite unreliable.

Taken in conjunction with the findings of previous studies, however, there would seem to be some support for the present theory. It is especially important to note, in this connection, that the predictions with respect to differential conditioning represent deductions from a theory developed on the basis of findings in a different situation, namely, simple classical and instrumental conditioning. In contrast to this sort of theory is the more particularistic hypothesis concerning anxiety and conditioned discrimination proposed by Hilgard, Jones, and Kaplan (10). These authors offered the conjecture that anxious persons would be more likely to perceive situations as threatening than do nonanxious persons, and, hence, as they put it, "a person whose discriminations are dimmed by anxiety" (10, p. 94) would be less likely to differentiate between the positive stimulus, which is threatening by virtue of its relation to the noxious air puff, and the negative stimulus, which is objectively nonthreatening.

In line with their hypothesis, Hilgard *et al.* found that the anxious *Ss* exhibited poorer discrimination than the nonanxious *Ss*. Although their anxious *Ss* gave

TABLE 5

Experiment II: Excitatory Strengths of Positive and Negative Stimuli on Trials 61–100 for Anxious and Nonanxious Men

Level of Drive	S_+		S_-		Diff.	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Anxious ($N=13$)	2.64	.78	2.00	.82	.64	.54
Nonanxious ($N=24$)	2.49	.71	1.94	.94	.54	.61

more CR's to the positive stimulus than did the nonanxious *Ss* both on Day 1, prior to the introduction of the differential conditioning procedure, and on Day 2, after differential conditioning was begun, the group differences were not significant on either day. In terms of percentage of CR's to the negative stimulus on Day 2, however, their anxious *Ss* were significantly higher. Further examination of these data, which were generously furnished by Hilgard, revealed that the difference in the percentage of responses on Day 2 to the positive and negative stimuli was greater for the nonanxious group than for their anxious group at about the .05 level on a single-tailed hypothesis (M's = 22.8% and 7.4%, respectively, $t = 1.72$, $df = 18$). Near the end of practice on Day 2 the direction of this result was reversed, the difference between S_+ and S_- being slightly and not significantly greater for their anxious *Ss*. However, their over-all results supported their hypothesis.

In contrast to these findings, the present studies showed that the difference between the E's to S_+ and S_- during differential conditioning was consistently greater for the anxious than for the nonanxious *Ss* in four independent instances. In view of their unreliability, these results do not permit the rejection of the null hypothesis with any great degree of certainty. On the other hand, they are even less favorable to the Hilgard, Jones, and Kaplan view than to our own. It does not appear possible, on the basis of our present information, to decide between these two very different theories.

Attention should be called to the point that no consideration has been given to the possible role of inhibition in our theoretical treatment of the present data. Undoubtedly, the effects of nonreinforcement played some role in the case of the response strength to the negative stimulus. It is possible that the failure to get consistent significant differences between the response strengths of the anxious and nonanxious *Ss* in the case of the negative stimulus, whereas significant differences were obtained in the case of the positive stimulus, is a function of this factor. Thus the occurrence of a greater number of responses on the part of the anxious *Ss* to the negative stimulus than on the part of the nonanxious *Ss* would lead to the development of a greater amount of inhibition in the anxious *Ss*. This, in turn, would have the effect of differentially reducing the response strength of the anxious *Ss*. A more clear-cut test of our present hypothesis would deal with the differential strength of response to a generalized stimulus uncomplicated by any possible differential effects of inhibition.

In conclusion, the findings of the present experiment suggest either that differences in anxiety produce a relatively small difference in D, or that the latter difference, if substantial in amount, is only a relatively minor variable contributing to the variance of conditioning scores. The failure to find consistent, significant differences in the present experiments also suggests that these anxiety differences play a relatively less important role in differential conditioning than in simple conditioning. In the latter situation the predicted differences between anxious and nonanxious groups have consistently been shown to be significant. Until we obtain a better understanding of what the many other factors determining conditioning performance are, and can thus effectively control them, it would appear necessary to employ relatively large numbers of *Ss* in our experiments.

Summary

Two experiments were conducted to provide data on the level of response and degree of differentiation between the positive and negative CS by anxious and nonanxious *Ss* in an eyelid conditioning situation. The positive CS in both experiments, a 500-cycle tone, and the negative CS, a 5,000-cycle tone, were each presented 50 times in a prearranged order. The interval between the positive CS and the UCS, a 1.0 lb./sq. in. air puff, was 500 msec. in both experiments. In Exp. I the UCS was omitted following the negative CS; in Exp. II, the negative CS was followed by the UCS at an interval of 2,500 msec.

The only significant differences between the anxious and nonanxious *Ss* occurred in connection with the level of response to the positive CS. No significant differences between the two

groups were found for the negative stimulus, although the level of response tended to be higher for the anxious *Ss*. Contrary to the findings reported by Hilgard, Jones, and Kaplan, the degree of differentiation between the positive and negative stimuli, as measured by the difference between their excitatory tendencies, was consistently greater for the anxious groups. None of the differences in degree of differentiation, however, was significant.

References

1. Amsel, A. The combination of a primary appetitional need with primary and secondary emotionally derived needs. *J. exp. Psychol.*, 1950, **40**, 1–14.
2. Brown, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, **33**, 209–226.
3. Brown, J. S., & Jacobs, A. The role of fear in the motivation and acquisition of responses. *J. exp. Psychol.*, 1949, **39**, 747–759.
4. Brown, J. S., Kalish, H. I., & Farber, I. E. Conditioned fear as revealed by magnitude of startle to an auditory stimulus. *J. exp. Psychol.*, 1951, **41**, 317–328.
5. Burke, C. J. Computation of the level of significance in the *F*-test. *Psychol. Bull.*, 1951, **48**, 392–397.
6. Grice, G. R., & Saltz, E. The generalization of an instrumental response to stimuli varying in the size dimension. *J. exp. Psychol.*, 1950, **40**, 702–708.
7. Heineman, C. E. A forced-choice form of the Taylor Anxiety Scale. *J. consult. Psychol.*, 1953, **17**, 447–455.
8. Hilgard, E. R., Campbell, A. A., & Sears, W. N. Conditioned discrimination: the development of discrimination with and without verbal report. *Amer. J. Psychol.*, 1937, **49**, 564–580.
9. Hilgard, E. R., Campbell, R. K., & Sears, W. N. Conditioned discrimination: the effect of knowledge of stimulus relationships. *Amer. J. Psychol.*, 1938, **51**, 498–506.
10. Hilgard, E. R., Jones, L. V., & Kaplan, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, **42**, 94–99.
11. Hovland, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, **17**, 125–148.
12. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
13. Hull, C. L. *A behavior system.* New Haven: Yale Univ. Press, 1952.
14. Kendler, H. H. Drive interaction: I. Learning as a function of the simultaneous presence of hunger and thirst drives. *J. exp. Psychol.*, 1945, **35**, 96–109.
15. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton Mifflin, 1953.
16. Margolius, G. Stimulus generalization of an instrumental response as a function of the number of reinforcements. *Amer. Psychologist*, 1953, **8**, 402. (Abstract)
17. McAllister, W. R. Eyelid conditioning as a function of the CS–US interval. *J. exp. Psychol.*, 1953, **45**, 417–422.
18. McAllister, W. R. The effect on eyelid conditioning of shifting the CS–US interval. *J. exp. Psychol.*, 1953, **45**, 423–428.
19. Miller, N. E. Studies of fear as an acquirable drive: I. Fear as motivation and fear reduction as reinforcement in the learning of new responses. *J. exp. Psychol.*, 1948, **38**, 89–101.
20. Mowrer, O. H. A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychol. Rev.*, 1939, **46**, 553–565.
21. Mueller, C. G. Numerical transformations in the analysis of experimental data. *Psychol. Bull.*, 1949, **46**, 198–223.
22. Perin, C. T. Behavior potentiality as a joint function of the amount of training and the degree of hunger at the time of extinction. *J. exp. Psychol.*, 1942, **30**, 93–113.
23. Skinner, B. F. *The behavior of organisms.* New York: D. Appleton-Century, 1938.
24. Spence, K. W. Mathematical formulations of learning phenomena. *Psychol. Rev.*, 1952, **59**, 152–160.
25. Spence, K. W., & Farber, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 116–119.
26. Spence, K. W., & Taylor, J. A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.
27. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
28. Warden, C. J. *Animal motivation studies.* New York: Columbia Univ. Press, 1931.
29. Webb, W. B. The motivational aspect of an irrelevant drive in the behavior of the white rat. *J. exp. Psychol.*, 1949, **39**, 1–14.

(Received for early publication October 16, 1953)

Journal of Experimental Psychology
Vol. 47, No. 2, 1954

ON READING PRINTED MATTER WITH INTERRUPTED LIGHT

VIRGINIA L. SENDERS

Antioch College

Gerathewohl and Taylor, in a paper recently published in this Journal (1), examined the possibility of using the Brücke-Bartley effect to increase the readability of low-contrast printed matter. They found that fewer lines of print on a chart of decreasing contrast could be read with an interrupted light than with a steady light *of the same intensity*. As a control for equality of light *energy*, they used lights interrupted for the same fraction of the cycle but at fairly high flash rates (50 flashes per second). Subjects were able to read about as well with the equal-energy control as with the flickering illumination. The authors therefore conclude that: "... the Bartley effect, if present, does not help to contrast the print on its background, because such an increase in brightness contrast would have improved necessarily the visibility of the text and thereby increased the number of lines read. What we found instead was a decrease in print visibility close to that produced by the 'Talbot effect'" (1, p. 281).

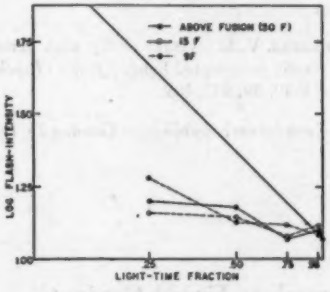


FIG. 1. The relationship between the logarithm of the flash intensity required to read a line of print and the logarithm of the light-time fraction. The 45° line is the contour for the constant Talbot (fused) brightness. Data are from Gerathewohl and Taylor (1, p. 281).

A re-examination of the data published by the above authors shows that the decrease in print visibility was not at all close to that produced by the Talbot effect. If we obtain, from the number of lines read, the luminous reflectance required and plot the logarithm of this number against the logarithm of the light time fraction (percentage of light in the cycle) we should find, according to Talbot's Law, a straight line with a slope of minus one. The results actually obtained are given in Fig. 1. The Talbot contour is also given for comparison.

What Gerathewohl and Taylor have actually shown is that, regardless of the flash frequency used, less energy is required for resolution when the light is interrupted than when it is steady. This is exactly the result I reported in this Journal in 1949 (2), and one set of my curves is reproduced in Fig. 2 for comparison with the Gerathewohl-Taylor data. That the decreased energy requirement does not result from the Brücke-Bartley effect is shown both by my data and by the data of Gerathewohl and Taylor.

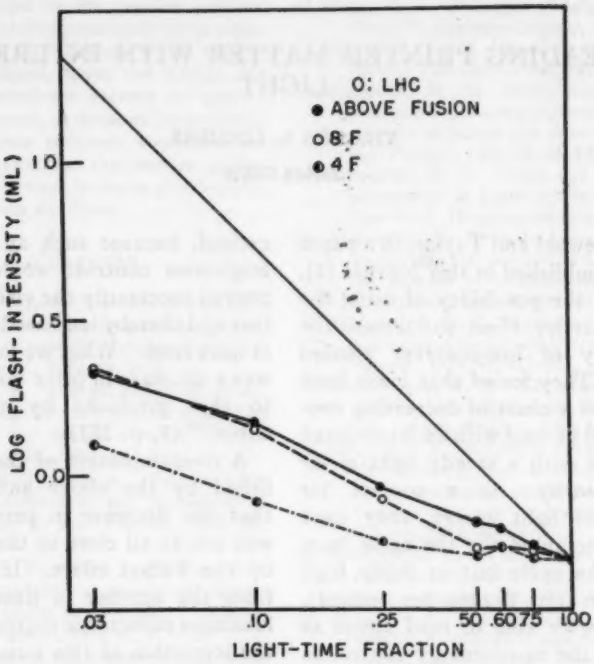


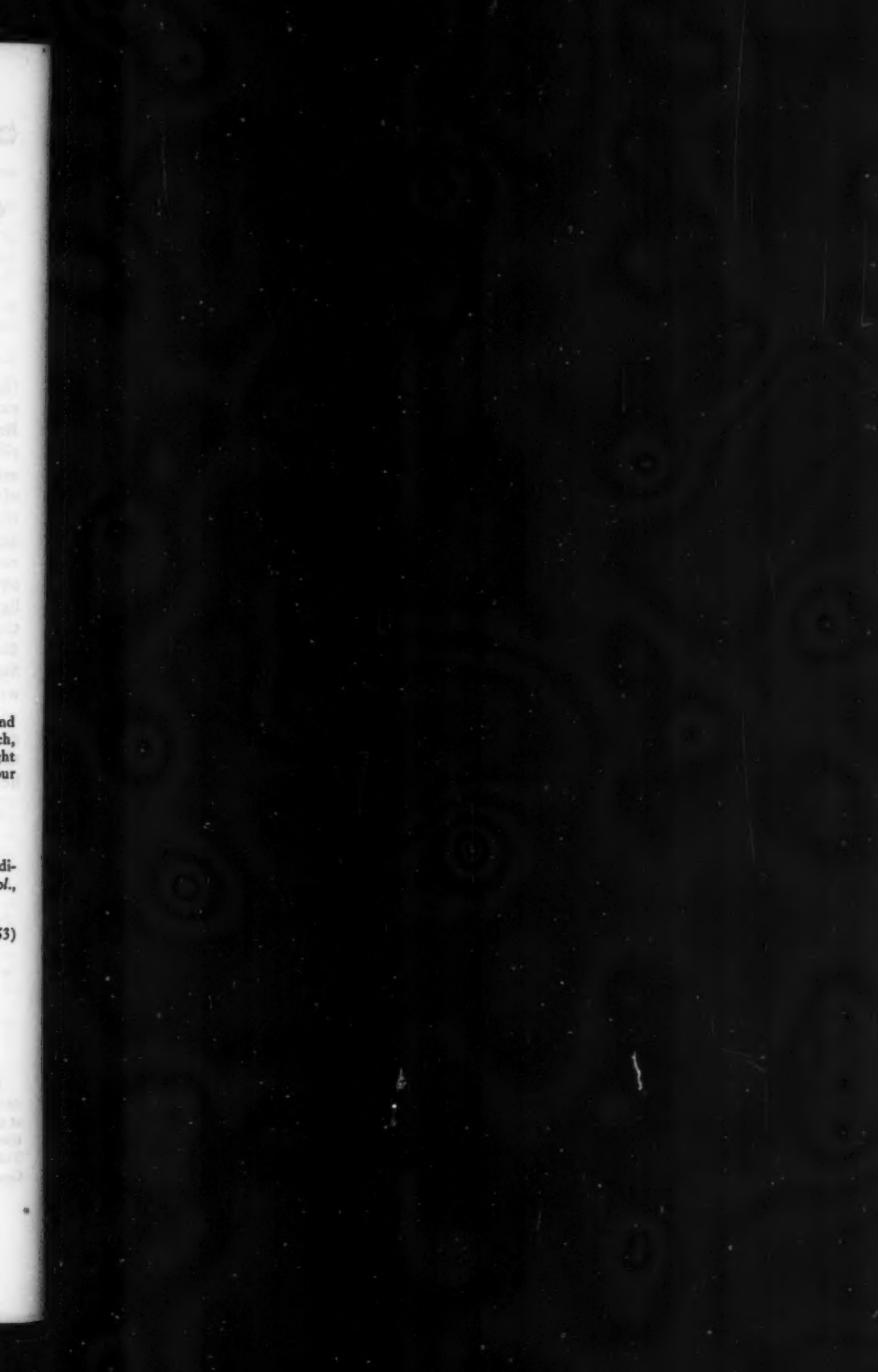
FIG. 2. The relationship between the logarithm of the flash intensity required for resolution and the logarithm of the light-time fraction, for different flash frequencies. A 2° × 2° test patch, centered at the fovea, is used with red light. All data are expressed as multiples of the steady light threshold for the series obtained with a flash frequency above fusion. The 45° line is the contour for the constant Talbot (fused) brightness. Data are from Senders (2, p. 456).

REFERENCES

1. GERATHEWOHL, S. J., & TAYLOR, W. F. Effect of intermittent light on the readability of printed matter under conditions of decreasing contrast. *J. exp. Psychol.*, 1953, **46**, 278–282.

2. SENDERS, V. L. Visual acuity with periodically interrupted light. *J. exp. Psychol.*, 1949, **39**, 453–465.

(Received for early publication October 29, 1953)

ERRATUM

In the article "A Quantitative Approach to Figural Goodness" by Julian Hochberg and Edward McAlister, which appeared in the November 1953 issue of this Journal (*J. exp. Psychol.*, **46**, 361–364) there was a typographical error in Table 1. The number of the "points of intersection" for "Cube Y" was printed as 17. It should have been 8.